

Grades 3 and 4
New Jersey Assessment of Skills and
Knowledge

TECHNICAL REPORT

March 2007

PTM 1506.58
Copyright © 2007 by New Jersey Department of Education
All rights reserved

STATE BOARD OF EDUCATION

RONALD K. BUTCHER Gloucester
President

ARCELIO APONTE Middlesex
Vice President

ARNOLD G. HYNDMAN Warren

KATHLEEN A. DIETZ Somerset

DEBRA ECKHERT-CASHA..... Morris

EDITHE FULTON..... Ocean

JOSEPHINE E. HERNANDEZ Union

FREDERICK J. LAGARDE, JR Passaic

ERNEST P. LEPORE Hudson

KENNETH J. PARKER..... Camden

DOROTHY S. STRICKLAND..... Essex

Lucille E. Davy, Commissioner
State Board of Education

TABLE OF CONTENTS

PART 1: INTRODUCTION	1
1.1 Description of the New Jersey Assessment of Skills and Knowledge (NJ ASK)	1
1.2 State-Level Results	2
1.3 NJ ASK Organizational Support.....	3
PART 2: TEST DEVELOPMENT	4
2.1 Test Specifications.....	4
Language Arts Literacy	5
Mathematics.....	5
Science.....	6
2.2 Development of Test Items.....	7
2.3 Item Review Process.....	8
2.4 Item Use.....	11
2.5 Test Forms Assembly	11
2.6 Quality Control for Test Construction.....	12
PART 3: TEST ADMINISTRATION	13
3.1 Participation.....	13
3.2 Test Security Procedures	14
3.3 Test Administration Procedures.....	14
3.4 Test Accommodations	15
PART 4: SCORING.....	16
4.1 Multiple Choice Items	16
4.2 Open Ended Items.....	16
Scorer Selection	17
Range Finding.....	17
Scorer Training	17
Scoring Procedures	18
4.3 Quality Control Procedures in Data Preparation	18
PART 5: STANDARD SETTING	19
5.1 Grade 4 Language Arts Literacy and Mathematics	19
5.2 Grade 3 Language Arts Literacy and Mathematics	20
5.3 Grade 4 Science	22
PART 6: TEST STATISTICS	25
6.1 Classical Item Statistics	25
6.2 Speededness.....	34
6.3 Intercorrelations.....	38
6.4 Item Bias Statistics.....	45
6.5 Summary Statistics	46

PART 7: SCALING AND EQUATING	55
7.1 Scaling	55
7.2 Equating Language Arts Literacy	56
7.3 Equating Mathematics	60
7.4 Equating Science.....	64
7.5 Rescore Equating Study.....	66
Part 8: VALIDITY	68
Content and Curricular Validity	68
PART 9: TEST RELIABILITY	69
9.1 Classical Reliability Estimates of the Test Scores.....	69
9.2 Reliability of Performance Classifications	73
9.3 Conditional Estimate of Error at Each Cut-Score.....	76
9.4 Rater Reliability	77
Part 10: REPORTING	79
10.1 Cycle I Reports	80
10.2 Cycle II Reports	85
10.3 State Summary Reporting.....	89
10.4 Interpreting Reports	89
10.5 Quality Control in Reporting	90
APPENDIX A: Statewide Cycle II Executive Summary Results	91
APPENDIX B: Additional Statewide Cycle II Results.....	103
APPENDIX C: Raw to Scale Score Conversions.....	126
REFERENCES.....	132

Tables and Figures

TABLE 1.2.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Total Student Group Tested – Grade 3	3
TABLE 1.2.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Total Student Group Tested – Grade 4	3
TABLE 2.1.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Total Points Possible by Content Area – Grades 3 & 4	7
Figure 2.3.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Item Approval Before Field Test.....	9
Figure 2.3.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Item Approval Before Operational Base Test	10
TABLE 2.3.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Number of Items Field Tested	11
TABLE 5.2.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Grade 3 LAL and Mathematics Standard-Setting Results from 2004 Recommended Cut Scores at the End of Phase 1 and Phase 2.....	22
TABLE 5.2.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Grade 3 LAL and Mathematics Standard-Setting Results from 2004 Recommended Cut Scores After Phase 2	22
TABLE 5.3.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Grade 4 Science Standard-Setting Results from 2005 Recommended Cut Scores at the End of Phase 1 and Phase 2	24
TABLE 5.3.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Grade 4 Science Standard-Setting Results from 2005 Recommended Cut Scores After Phase 2.....	25
TABLE 6.1.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Item Difficulty and Discrimination Summary Statistics for Dichotomously Scored and Open-Ended Items by Test Section and Cluster – Grade 3	27
TABLE 6.1.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Item Difficulty and Discrimination Summary Statistics for Dichotomously Scored and Open-Ended Items by Test Section and Cluster – Grade 4.....	28
TABLE 6.1.3 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Frequency Distributions of Item Difficulty Values and Biserial Discrimination Indices by Content Cluster Language Arts Literacy – Grade 3	29
TABLE 6.1.4 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Frequency Distributions of Item Difficulty Values and Biserial Discrimination Indices by Content Cluster Language Arts Literacy – Grade 4	30

TABLE 6.1.5 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Frequency Distributions of Item Difficulty Values and Biserial Discrimination Indices by Content Cluster Mathematics – Grade 3	31
TABLE 6.1.6 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Frequency Distributions of Item Difficulty Values and Biserial Discrimination Indices by Content Cluster Mathematics – Grade 4	32
TABLE 6.1.7 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Frequency Distributions of Item Difficulty Values and Biserial Discrimination Indices by Content Cluster Science – Grade 4.....	33
TABLE 6.2.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Percentage of Students Omitting the Last Items of Each Test Part – Grade 3	35
TABLE 6.2.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Percentage of Students Omitting the Last Items of Each Test Part – Grade 4	36
TABLE 6.3.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Intercorrelations Among Major Content Clusters and Item Types – Grade 3	38
TABLE 6.3.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Intercorrelations Among Major Content Clusters and Item Types – Grade 4.....	39
TABLE 6.3.3 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Intercorrelations Among Content Areas and Clusters – Grade 3	41
TABLE 6.3.4 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Intercorrelations Among Content Areas and Clusters – Grade 4	42
TABLE 6.4.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) DIF Categories	46
TABLE 6.5.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Means and Standard Deviations of Students’ Raw Scores by Test Section – Grade 3.....	47
TABLE 6.5.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Means and Standard Deviations of Students’ Raw Scores by Test Section – Grade 4.....	47
TABLE 6.5.3 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Means and Standard Deviations of Students’ Raw Scores and Percent Correct by Content Area – Grade 3	49
TABLE 6.5.4 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Means and Standard Deviations of Students’ Raw Scores and Percent Correct by Content Area – Grade 4	50
TABLE 6.5.5 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Means and Standard Deviations of Students’ Raw Scores and Percent Correct on the Dichotomously Scored Items by Content Area – Grade 3.....	51
TABLE 6.5.6 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Means and Standard Deviations of Students’ Raw Scores and Percent Correct on the Dichotomously Scored Items by Content Area –Grade 4.....	52

TABLE 6.5.7 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Means and Standard Deviations of Students' Raw Scores and Percent Correct on the Open-Ended Items by Content Areas and Clusters – Grade 3	53
TABLE 6.5.8 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Means and Standard Deviations of Students' Raw Scores and Percent Correct on the Open-Ended Items by Content Areas and Clusters – Grade 4	54
TABLE 7.1.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Scaling Parameters for Base Forms.....	56
Figure 7.2.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Language Arts Literacy Backwards and Forward Equating Designs.....	57
TABLE 7.2.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Language Arts Literacy Item Parameters – Grade 3	58
TABLE 7.2.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Language Arts Literacy Fixed OE Item Step Parameters – Grade 3.....	59
TABLE 7.2.3 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Language Arts Literacy Item Parameters – Grade 4	59
TABLE 7.2.4 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Language Arts Literacy Fixed OE Item Step Parameters – Grade 4.....	60
TABLE 7.3.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Mathematics Item Parameters – Grade 3	61
TABLE 7.3.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Mathematics OE Item Step Parameters – Grade 3	62
TABLE 7.3.3 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Mathematics Item Parameters – Grade 4	63
TABLE 7.3.4 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Mathematics OE Item Step Parameters – Grade 4	64
TABLE 7.4.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Science OE Item Step Parameters – Grade 4	66
TABLE 9.1.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Reliability Estimates and Standard Errors of Measurement (SEM) for Content Areas and Clusters – Grade 3	70
TABLE 9.1.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Reliability Estimates and Standard Errors of Measurement (SEM) for Content Areas and Clusters – Grade 4.....	71
TABLE 9.1.3 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Reliability Estimates and Standard Errors of Measurement (SEM) for Dichotomously Scored Items Within Content Clusters – Grade 3	72

TABLE 9.1.4 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Reliability Estimates and Standard Errors of Measurement (SEM) for Dichotomously Scored Items Within Content Clusters – Grade 4.....	72
TABLE 9.2.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Reliability of Classification and Decision Accuracy – Grade 3	74
TABLE 9.2.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Reliability of Classification and Decision Accuracy – Grade 4.....	75
TABLE 9.3.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Conditional Estimate of Error at Each Cut-Score	77
TABLE 9.4.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Consistency Between Raters Scoring Writing Tasks and Open-Ended Items - Grade 3.....	78
TABLE 9.4.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Consistency Between Raters Scoring Writing Tasks and Open-Ended Items – Grade 4.....	79
Figure 10.1.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Individual Student Report (ISR) – Front	81
Figure 10.1.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) Individual Student Report (ISR) – Back.....	82
Figure 10.2.1 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) School Report - Performance by Demographic Groups.....	86
Figure 10.2.2 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) School Report - Summary of School Cluster Performance	87
TABLE A.3.1 STATEWIDE PERFORMANCE BY DEMOGRAPHIC GROUPS SPRING 2007 -- GRADE 3 Language Arts Literacy	94
TABLE A.3.2 STATEWIDE PERFORMANCE BY DEMOGRAPHIC GROUPS SPRING 2007 -- GRADE 3 Mathematics	95
TABLE A.4.1 STATEWIDE PERFORMANCE BY DEMOGRAPHIC GROUPS SPRING 2007 -- GRADE 4 Language Arts Literacy	100
TABLE A.4.2 STATEWIDE PERFORMANCE BY DEMOGRAPHIC GROUPS SPRING 2007 -- GRADE 4 Mathematics	101
TABLE A.4.3 STATEWIDE PERFORMANCE BY DEMOGRAPHIC GROUPS SPRING 2007 -- GRADE 4 Science.....	102
TABLE B.3.1 NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE BY DISTRICT FACTOR GROUP LANGUAGE ARTS LITERACY SECTION – Grade 3	103

TABLE B.3.2 NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE BY DISTRICT FACTOR GROUP MATHEMATICS SECTION – Grade 3	106
TABLE B.3.3 NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE LANGUAGE ARTS LITERACY AND MATHEMATICS FOR THE SPECIAL NEEDS DISTRICTS AS COMPARED TO ALL OTHER DISTRICTS – Grade 3	109
TABLE B.4.1 NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE BY DISTRICT FACTOR GROUP LANGUAGE ARTS LITERACY SECTION – Grade 4	111
TABLE B.4.2 NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE BY DISTRICT FACTOR GROUP MATHEMATICS SECTION – Grade 4	114
TABLE B.4.3 NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE BY DISTRICT FACTOR GROUP SCIENCE SECTION – Grade 4	117
TABLE B.4.4 NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE LANGUAGE ARTS LITERACY, MATHEMATICS, AND SCIENCE FOR THE SPECIAL NEEDS DISTRICTS AS COMPARED TO ALL OTHER DISTRICTS – Grade 4	120
TABLE C.3.1 Conversion 2007 NJ ASK Language Arts Literacy – Grade 3	127
TABLE C.3.2 Conversion 2007 NJ ASK Mathematics – Grade 3	128
TABLE C.4.1 Conversion 2007 NJ ASK Language Arts Literacy – Grade 4	129
TABLE C.4.2 Conversion 2007 NJ ASK Mathematics – Grade 4	130
TABLE C.4.3 Conversion 2007 NJ ASK Science – Grade 4	131

PART 1: INTRODUCTION

The purpose of this Technical Report is to provide information about the New Jersey Assessment of Skills and Knowledge (NJ ASK) administered as an operational assessment in March 2007. This report is intended for use by those who evaluate tests, interpret scores, or use test results for making educational decisions. It includes the following sections: test development, test administration, scoring, standard setting, item level statistics, scaling and equating, test statistics, validity, and score reporting. It includes references to additional reports and documents available for the NJ ASK.

1.1 Description of the New Jersey Assessment of Skills and Knowledge (NJ ASK)

The spring 2007 New Jersey Assessment of Knowledge and Skills (NJ ASK) was administered to students in grades three and four. It consisted of two content areas in grade 3, Language Arts Literacy and Mathematics, and three content areas in grade 4, Language Arts Literacy, Mathematics, and Science. Science was administered as an operational test to grade four students for the first time in 2005. The NJ ASK is designed to give an early indication of the progress students are making in mastering the knowledge and skills described in the Core Curriculum Content Standards. The results are to be used by schools and districts to help identify strengths and weaknesses in their educational programs. It is anticipated that this process will lead to improved instruction and better alignment with the Core Curriculum Content Standards in kindergarten through grade four. The results may also be used, along with other indicators of student progress, to identify those students who may need instructional support in any of the content areas. This support, which could be in the form of individual or programmatic intervention, would be a means to address any identified knowledge or skill gaps.

The NJ ASK scores are reported as scale scores and performance levels in each of the content areas. Following are the score ranges and their associated performance level.

- 100-199 Partially Proficient
- 200-249 Proficient
- 250-300 Advanced Proficient

The scores of students who are included in the Partially Proficient level are considered to be below the state minimum of proficiency and those students may be in need of instructional support.

The NJ ASK was administered from March 19 through March 23, 2007. The 2007 Language Arts Literacy and Mathematics tests were administered to 102,812 total students in grade 3. Performance levels for the grade 3 NJ ASK tests were established by panels of educators during sessions held in June, 2004 and were approved by the New Jersey State Board of Education on July 7, 2004. The 2007 Language Arts Literacy, Mathematics, and Science tests were administered to 102,490 total students in grade 4. The grade 4 performance standards for Mathematics were set in 1999 and the standards for grade 4 Language Arts Literacy were established in 2001. Performance levels for the grade 4 NJ ASK Science test was established by

a panel of educators during sessions held in June, 2005 and performance standards were approved by the New Jersey State Board of Education on July 6, 2005.

1.2 State-Level Results

This section includes two tables summarizing statewide test results for the 2007 administration of the NJ ASK. Tables 1.2.1 and 1.2.2 show the number and percentage of students in each performance category (i.e., Partially Proficient, Proficient, and Advanced Proficient) for each subject in grades 3 and 4, respectively. The “number of students tested” is based on all students who received a test booklet, excluding those who were voided, not present or APA exempt with no scale scores.

NOTE: Percentages shown in tables through this *Technical Report* may not total 100 due to rounding.

Following is a list of five state-level highlights for all students.

- Of the 100,877 grade 3 students with valid scale scores in Language Arts Literacy in spring 2007, 16.6% scored in Partially Proficient; 75.2% scored in Proficient and 8.2% scored in Advanced Proficient (Table 1.2.1).
- Of the 101,800 grade 3 students with valid scale scores in Mathematics in spring 2007, 12.7% scored in Partially Proficient; 55.0% scored in Proficient and 32.3% scored in Advanced Proficient (Table 1.2.1).
- Of the 100,617 grade 4 students with valid scale scores in Language Arts Literacy in spring 2007, 19.4% scored in Partially Proficient; 73.9% scored in Proficient and 6.7% scored in Advanced Proficient (Table 1.2.2).
- Of the 101,310 grade 4 students with valid scale scores in Mathematics in spring 2007, 15.3% scored in Partially Proficient; 43.7% scored in Proficient and 41.0% scored in Advanced Proficient (Table 1.2.2).
- Of the 101,266 grade 4 students with valid scale scores in Science in spring 2007, 16.9% scored in Partially Proficient; 42.3% scored in Proficient and 40.8% scored in Advanced Proficient (Table 1.2.2).

TABLE 1.2.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Total Student Group Tested – Grade 3**

TEST SECTION	NUMBER ^a OF VALID SCALE SCORES	PROFICIENCY LEVELS						SCALE SCORE MEAN
		PARTIALLY PROFICIENT (100-199)		PROFICIENT (200-249)		ADVANCED PROFICIENT (250-300)		
		No.	%	No.	%	No.	%	
LANGUAGE ARTS LITERACY 2007	100,877	16,760	16.6%	75,893	75.2%	8,224	8.2%	218.0
MATHEMATICS 2007	101,800	12,887	12.7%	56,037	55.0%	32,876	32.3%	232.3

a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT, AND APA EXEMPT WITH NO SCALED SCORES.

TABLE 1.2.2

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Total Student Group Tested – Grade 4**

TEST SECTION	NUMBER ^a OF VALID SCALE SCORES	PROFICIENCY LEVELS						SCALE SCORE MEAN
		PARTIALLY PROFICIENT (100-199)		PROFICIENT (200-249)		ADVANCED PROFICIENT (250-300)		
		No.	%	No.	%	No.	%	
LANGUAGE ARTS LITERACY 2007	100,617	19,525	19.4%	74,371	73.9%	6,721	6.7%	215.7
MATHEMATICS 2007	101,310	15,525	15.3%	44,238	43.7%	41,547	41.0%	234.1
SCIENCE 2007	101,266	17,071	16.9%	42,885	42.3%	41,310	40.8%	231.3

a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT, AND APA EXEMPT WITH NO SCALED SCORES.

1.3 NJ ASK Organizational Support

The NJ ASK is administered by the Office of State Assessments within the Department of Education. The staff of the Office of State Assessments directs the implementation of the statewide assessment programs. In addition to planning, scheduling, and directing all NJ ASK

activities, the staff is extensively involved in numerous test review, security, and quality control procedures.

In 2003, the contract for developing and administering the NJ ASK was awarded to Educational Testing Service (ETS). ETS is the primary contractor working in partnership with Pearson to provide the full range of testing services for the NJ ASK. ETS activities include program management, item development, test form development, publication development and printing, customer service and Web site hosting, and supporting regional workshops for district and school test coordinators. ETS also provides research and psychometric services in support of the NJ ASK statewide assessment program. The major activities supported by Pearson include: printing test books; distributing assessment materials in a secure manner; receiving, scanning, editing and scoring the answer documents; packaging, transporting and scoring open-ended responses; and score reporting.

PART 2: TEST DEVELOPMENT

The Elementary School Proficiency Assessment (ESPA) was first administered as an operational test at grade 4 from 1999 through 2002 to provide an early indication of student progress toward achieving the knowledge and skills identified in the Core Curriculum Content Standards (CCCS). ESPA was replaced in spring 2003 with the New Jersey Assessment of Skills and Knowledge (NJ ASK), a comprehensive, multi-grade assessment program. The purpose of these assessments is to provide indicators of student progress and to identify students who need additional instructional support in order to reach the CCCS. Details of the NJ ASK test development process are presented in this section.

2.1 Test Specifications

During the summer of 1996, three content committees consisting of 46 New Jersey educators developed the Elementary School Proficiency Assessment Content Domain Outline (February 1997), and a directory of test specifications and sample items for each content area to provide content/skill outlines and sample items. These directories describe the test, format of the items, and the scores to be generated by the test. This test specification work done by New Jersey educators serves as the foundation for all test item development.

The committees of New Jersey educators rely upon their expertise and the New Jersey Core Curriculum Content Standards to design a test that is universally accessible to all grade 3 and grade 4 students and is composed of test questions that are age- and grade-appropriate. The test specifications, released sample assessments and holistic scoring guides are designed for use by curriculum specialists and teachers to improve instruction at the district, school and classroom levels.

In 2003, the ESPA became the NJ ASK. The NJ ASK is designed to measure the same Core Curriculum Content Standards as the ESPA. The items and test format of the NJ ASK are similar

to those of the ESPA. Brief descriptions of the test content measured with the NJ ASK are presented in the following sections.

Language Arts Literacy

The Language Arts Literacy section of each test measures students' achievements in reading and writing. Students read passages selected from published books, newspapers, and magazines as well as everyday text, and respond to related multiple-choice and open-ended questions.

The Language Arts Literacy assessment currently assesses knowledge and skills in the following clusters (A "cluster" is a group of related test questions on a single topic):

- Writing
 - Writing about Pictures
 - Writing About Poems
- Reading
 - Working with Text
 - Analyzing Text

For an in depth description of the NJ ASK Language Arts Literacy including specifications, visit the NJ Department of Education website at:

http://www.nj.gov/education/njpep/assessment/njask_lal/Overview_njask_lal.pdf

Mathematics

The Mathematics section of each test measures students' ability to solve problems by applying mathematical concepts. The NJ ASK assesses four Core Curriculum Content Standards in Mathematics:

- Number Sense and Numerical Operations
- Geometry and Measurement
- Patterns and Algebra
- Data Analysis, Probability, and Discrete Mathematics

A process cluster, Problem Solving, is also reported on score reports. The process cluster refers to test questions that measure mathematical problem-solving ability. Each test question on the Mathematics assessment measures one content cluster and may contribute to the process cluster. Each cluster in Mathematics contains one open-ended item except for Number and Numerical Operations, which contains two. For an in-depth description of the NJ ASK Mathematics Test Specifications visit the NJ Department of Education website at:

<http://www.nj.gov/education/njpep/assessment/TestSpecs/MathNJASK/index.html>

Science

The Science section measures students' ability to recall information and to solve problems by applying science concepts. The NJ ASK assesses 10 core curriculum content standards – with a focus on the Life, Physical, and Earth clusters. The standards for Science are:

- Scientific Processes
- Science and Society
- Mathematical Applications
- Nature and Process of Technology
- Characteristics of Life
- Chemistry
- Physics
- Earth Science
- Astronomy and Space Science
- Environmental Studies

The reported Science clusters are Life Science, Physical Science, Earth Science and Application. For an in-depth description of the NJ ASK Science Test Specifications visit the NJ Department of Education website at:

<http://www.nj.gov/education/njep/assessment/TestSpecs/ScienceNJASK/index.html>

Table 2.1.1 summarizes the total points possible for each of the content areas of the operational NJ ASK administered in March 2007 for grades 3 and 4.

TABLE 2.1.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Total Points Possible by Content Area – Grades 3 & 4**

Language Arts Literacy	Grade 3	Grade 4
Total	40 points	43 points
Writing	20 points	20 points
Writing/Picture	10 points	10 points
Writing/Poem	10 points	10 points
-----	-----	-----
Reading	20 points	23 points
Working with Text	8 points	7 points
Analyzing Text	12 points	16 points
Mathematics	Grade 3	Grade 4
Total	33 points	43 points
Number Sense & Numerical Operations	9 points	13 points
Geometry & Measurement	8 points	10 points
Patterns & Algebra	8 points	10 points
Data Analysis, Probability & Discrete Math	8 points	10 points
-----	-----	-----
Problem Solving	12 points	23 points
Science		Grade 4
Total		39 points
Life Science		15 points
Physical Science		12 points
Earth Science		12 points
Application		33 points

* Within a content area, cluster-level results show how students perform on the sets of items that measure particular knowledge and skills (clusters above the dotted line) or particular processes (clusters below the dotted line). Though an item on the NJ ASK can contribute to a cluster above the line (for example, Reading) as well as a cluster below the line (for example, Working with Text), each item is counted only once in the total score.

2.2 Development of Test Items

The NJ ASK consists of two types of items:

1. Operational or base test items used to determine students' scores and
2. Field-test items evaluated for use as future base test items.

A team of Educational Testing Service (ETS) subject area specialists and consulting item writers begin the NJ ASK item development process. These writers are teachers or former teachers who have a great deal of specialized knowledge concerning their area of content expertise. All item writers for the NJ ASK program receive in-depth training about the NJ ASK and item writing practice and standards.

The following steps outline the item development process:

1. NJDOE and ETS: Review test specification and create item specifications
2. ETS: Select and train item writers
3. Item Writers: Write test items
4. ETS: Conduct initial item review
5. ETS: Conduct item review by experienced senior staff
6. NJDOE: Conduct content and bias review with test committees

7. Items are field tested
8. NJDOE: Conduct Statistical Item Review with test committees
9. Approved items go into the item bank

The ETS item development process for each testing cycle begins with a formal review of the Core Curriculum Content Standards and the item specifications. The NJ ASK Item Specifications detail the standards to be measured, the number of items to be written, the item formats to be used, and other specific directions for developing the items. All NJ ASK items must be written to measure the New Jersey Core Curriculum Content Standards.

Item-writer training sessions are convened by content area at a variety of locations. These training sessions provide guidelines for the item writing process and demonstrate how to write items that are free of bias and sensitivity issues. Each consulting item writer is asked to sign a Letter of Agreement. This letter specifies the confidentiality and security regulations and outlines the ownership regulations. At the start of the item writing process, each item writer is provided with the following materials:

- An overview of the New Jersey Assessment of Skills and Knowledge
- Test specifications for each subject area and item specifications
- A description of the item formats to be used, including important characteristics of each format
- A description of the item writing process and measures to avoid writing biased items
- A listing of the security procedures to be followed during the item development process

All items written by item writers are reviewed, revised, and edited by ETS subject area specialists and editors prior to review by the New Jersey Test Committees. Before any item is included on a field test or operational base test, it must have the approval of the committees, as well as the NJDOE.

As items are developed, ETS documents each item's relevancy to the Core Curriculum Content Standards and the directories of test specifications. During this process, each item is assigned a unique item identification number. The number is used to track the item throughout the development process and later in the item bank.

2.3 Item Review Process

Once test items have been through initial item review and review by experienced senior staff at ETS, they are prepared for test committees' reviews. Before any item is included on a field test or operational test, it must have the approval of the New Jersey Assessment Content and Sensitivity Review Committees. Typically, the committees consist of experienced educators and curriculum experts. Committee members also represent the diversity of the state in terms of ethnicity and geographic regions.

The New Jersey Test Committee members provide expert judgments as to the alignment of each test item with the Core Curriculum Content Standards and the content-specific test specifications. Committee members are selected based on their level of content area knowledge

and number of years of teaching experience. Additionally, special care is taken to select members who are representative of the various districts and District Factor Groups (DFGs) within the State. Prior to field testing, the Office of State Assessments staff and the Language Arts Literacy, Mathematics, or Science Committees review all items. The Committees review each test item to determine if the item meets test specifications and address an appropriate level of difficulty. Committees also ensure that test items are not offensive and do not reinforce negative stereotypes, and that test items appropriately reflect multicultural society. Figure 2.3.1 presents a sample of the form that must be marked “Definitely Use” or “Revise and Use With Approval” during review committee meetings before an item is included on a field test.

Figure 2.3.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Item Approval Before Field Test**

Sensitivity		Content		
*Comments		*Comments		
Sensitivity Issue	Yes No	Meets Specifications	Yes	No
If Yes, identify category and explain*		Appropriate Difficulty	Yes	No
		Accurate Coding	Yes	No
Definitely Use		Definitely Use		
Revise and Use With Approval		Revise and Use With Approval		
Revise and Resubmit		Revise and Resubmit		
Do Not Use*		Do Not Use*		

Sensitivity Sign-off _____ Date _____ Content Chairperson’s Signature _____ Date _____

All test items are field tested and reviewed again before they can be used as operational test items. The committees meet to review the item statistics, which include: item means, response frequencies, biserial correlations (with operational test total scores), and other descriptive statistics. Prior to the presentation of items and statistics to reviewers, the New Jersey Department of Education defined boundaries within which item statistics should fall to be considered usable for future forms. In general, items with p-values below .30 or above 0.90 were considered to be usable only if a strong content argument could be made for their inclusion in the item bank. An item could be flagged for low or high p-value and/or low biserial correlation with base test total scores.

Also, for the statistical item review, the Mantel-Haenszel statistic is calculated to show whether or not students are responding to an item in a way that their overall ability (as measured by the base test) would lead us to expect. The statistic allows the committees to examine group membership (by ethnicity or by gender). The Mantel-Haenszel statistic is used for a classification determination of category A, B, or C. An item in Category A shows no or minor relationship between group membership and performance. Category B items show small to moderate

relationship between membership and performance. Category C items show a substantial relationship between group membership and item performance and must be examined carefully by the committees to make sure these items are not biased.

Figure 2.3.2 presents a sample of the form that must be marked “Definitely Use” or “Revise and Use With Approval” during review committee meetings of the field-test statistics before an item is included on an operational base test.

Figure 2.3.2

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Item Approval Before Operational Base Test**

Sensitivity			Content		
*Comments			*Comments		
Sensitivity Issue	Yes	No	Appropriate Difficulty	Yes	No
If Yes, identify category and explain*			P-Value = 0.65		
Mantel-Haenszel Category C			Biserial = 0.42		
W-AA _____	W-H _____	M-F _____			
Definitely Use			Definitely Use		
Revise and Use With Approval			Revise and Use With Approval		
Revise and Resubmit			Revise and Resubmit		
Do Not Use*			Do Not Use*		

Sensitivity Sign-off _____ Date _____ Content Chairperson's Signature _____ Date _____

Table 2.3.1 shows the number of field-test items presented during the March 2007 field-test administration. A sampling plan was developed that randomly assigned field-test forms to districts. To the extent possible, this plan ensured that the student group taking each field-test form would be representative of the DFG distribution of the New Jersey districts.

TABLE 2.3.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Number of Items Field Tested**

		Multiple-Choice Items		Open-Ended Items		Writing Activities	
		Presented	Accepted	Presented	Accepted	Presented	Accepted
Grade 3	Language Arts	60	52	16	15	4	4
	Mathematics	70	59	10	9	--	--
		Presented	Accepted	Presented	Accepted	Presented	Accepted
Grade 4	Language Arts	54	47	20	20	6	6
	Mathematics	60	54	20	16	--	--
	Science	200	162	17	10	--	--

2.4 Item Use

All field-test items approved for use on an operational test form are moved into the item bank. Test development staff members choose from the available banked items when building an operational test form. In most cases, a test item is used operationally one time, unless the item is used a second time as an anchor item. After operational use, items are retired. A small number of previously used items have been released for practice.

2.5 Test Forms Assembly

There are four steps associated with assembling test forms for NJ ASK:

1. Determine form design
 2. Select items that meet content specifications
 3. Evaluate statistical specifications and select items to meet these specifications
 4. Review and approve test forms
- 1) **Determine forms design** – Each form consists of a set of operational items plus a set of variable items. The variable items provide opportunities for meeting equating needs and field-testing new items. The number of variable sections for each grade and subject is dependent upon the pool of items available for field-testing.
- 2) **Select items that meet content specifications** – Each content area measures subsets of items called clusters. In Language Arts Literacy the clusters include: Writing (Writing about Pictures and Writing about Poems), and Reading (Working with Text and Analyzing Text).

In Mathematics the clusters include: Number Sense and Numerical Operations; Geometry and Measurement; Patterns and Algebra; and Data Analysis, Probability, and Discrete Mathematics. There is also a process cluster called Problem Solving. In Science the clusters include: Life Science, Physical Science and Earth Science. Science also has a process cluster called Application. Test forms must be similar to previous NJ ASK forms in terms of the number of items, the number of points, and the distribution of the content.

- 3) **Evaluate statistical specifications** – As forms are created it is necessary to determine if the statistical specifications have been met. Statistical specifications based on previous forms provide guidelines for building new test forms. Spreadsheets (form matrices) are used to provide information on the statistical properties of newly created forms. These matrices contain the following statistics: Average p-value, biserial correlation and average IRT difficulty (among other statistics). These data are reviewed to make certain that current forms are not substantially harder or easier than previous forms. Linking designs are also evaluated at this stage.
- 4) **Final approval of forms** – Once the content and statistical specifications have been met for each grade and subject, the forms are approved by the ETS Statistical Coordinator and by the NJ DOE. The forms are then released for production and editorial reviews.

Checklists and quality control procedures accompany each stage of form development. Some of these procedures are listed below:

2.6 Quality Control for Test Construction

Following is a list of quality control procedures used during the assembly of NJ ASK forms:

- Construct forms based on all content requirements noted in the test specifications.
- Verify correct number of items per standard or reporting category based on test specifications.
- Review selected items to ensure a wide sampling of the knowledge and skills being measured.
- Ensure that all selected items have been through the appropriate review procedures and are approved for use by the NJ DOE.
- Check for a variety of item topics, equal distribution of male/female, ethnicities, etc.
- Verify appropriate portions of items with and without artwork.
- Check for cueing across all items on each form.
- Verify match of unique item identification numbers (UIN) to test matrix.
- Verify equal or nearly equal distribution of answer choices for MC items.
- Verify and document items needing manipulative sheets (Mathematics only).
- Ensure that the test meets the statistical specifications.
- Verify match of statistical data on item card to statistical data on test matrix.
- Consider any statistical flags or problems.
- Check statistics to ensure that the collection of items yields an overall difficulty that falls within the specified range.

- ❑ Verify that items have not been released to the public.
- ❑ Verify equal or nearly equal distribution of answer choices for MC items.
- ❑ Verify correct answer key for each item.
- ❑ Content review of form by senior staff.
- ❑ Statistical review of form by psychometrician.
- ❑ Send form to NJ DOE for review and approval.

PART 3: TEST ADMINISTRATION

The Spring 2007 New Jersey Assessment of Skills and Knowledge (NJ ASK) included Grade 3 and Grade 4 testing sections in Language Arts Literacy and Mathematics as well as Science in Grade 4. The Language Arts Literacy section consists of reading passages, multiple-choice items, open-ended items, and writing tasks. The Language Arts Literacy section is administered over two days for both grades. The Mathematics section consists of multiple-choice and open-ended items that must be answered with the use of a calculator, and multiple-choice items that must be answered without the use of a calculator. The Mathematics section is administered over a two-day period for Grade 4 and a one-day period for Grade 3. The Science section, which consists of multiple-choice and open-ended items, is administered on one day.

Field-test items for all tests are embedded within the sections of the regular test. The make-up tests are scheduled by school districts for administration any morning during the week following the regular NJ ASK administration. Districts have the flexibility to choose which subjects are tested on which days of the make-up period.

3.1 Participation

General Education Students

The NJ ASK must be administered to all third- and fourth-grade students in New Jersey public schools except those whose Individual Education Program exempts them from taking the NJ ASK.

Limited English Proficient Students

Limited English Proficient (LEP) students must take the test according to federal guidelines for the No Child Left Behind (NCLB) Act of 2001.

Students with Disabilities

Students with Disabilities in the third- and fourth-grade eligible for special education under the Individuals with Disabilities Education Act or eligible under Section 504 of the Rehabilitation Act of 1973 must take each subject area of the NJ ASK unless their Individualized Education Program (IEP) or 504 plan specifically states that they will not participate in one or more subject areas of the test. Students who are ungraded must take the NJ ASK in the calendar year in which

they are 9, 10, or 11 years old and when they are first instructed in the knowledge and skills tested. Students whose IEP exempts them from participation in the NJ ASK must participate in the Alternate Proficiency Assessment (APA).

3.2 Test Security Procedures

Standard Security Procedures

The NJ ASK test booklets and their contents are secure materials. Detailed procedures for maintaining the security of test materials while test materials are in the districts are outlined in the Test Administration Manual. It is the responsibility of school districts to guarantee the security of the test materials. Examiners, proctors, and other school personnel are prohibited from copying, reading, discussing, or disclosing any test items before, during, or after the test administration. When not being used during a test period, test materials are stored in a secure, locked place that is accessible only to individuals whose access is authorized by the school test coordinator. Inventory forms track test materials as they move from one location to another within the districts.

Security Breach Procedures

Breach test forms and examiner manuals are prepared in the event of a security breach. If the New Jersey Department of Education (NJ DOE) identifies a security breach during the test administration window, the subcontractor removes the NJ ASK test materials from the involved district or school. The test books for the subject area affected are coded with a void code 5 indicating a security breach. If time permits (determined by NJ DOE) breach forms are delivered to the districts and districts are required to test the affected students in the subject area impacted. When students are re-tested during the test administration window, scores are reported based on the breach form test scores. If a security breach is identified after the test administration window, the impacted test books are coded void code 5 (security breach) and no test results are reported for that subject area. Students receive a score for the subject area that was not impacted by the security breach.

3.3 Test Administration Procedures

School test coordinators, examiners and proctors are responsible for the administration of the exam. Their responsibilities include

- distributing test materials each morning of testing,
- overseeing the recording on School Security Checklists of the transfer of test booklets,
- supervising testing, ensuring proper test administration procedures are followed according to the instructions in the provided Examiner Manuals,
- ensuring that accommodations/modifications listed in the IEPs/504 plans of students with disabilities are implemented,
- monitoring any potential circumstances that may seriously interrupt or interfere with the test administration,

- reporting any testing irregularities that occur during the administration,
- notifying district test coordinator immediately of any missing test booklets,
- scheduling make-up testing for any students who missed one or more days of the regular testing period, and
- returning testing materials to contractors.

3.4 Test Accommodations

General Education Students

General education students receive no special testing accommodations other than the standard room setup and materials distribution described in the Examiner Manual.

Accommodations and Modifications for Students with Disabilities

To ensure that students are tested under appropriate conditions, the Department of Education has adopted test accommodations and modifications that may be used when testing special populations of students. The content of the test typically remains the same, but administration procedures, setting, and answer modes may be adapted. Students requiring accommodations must be tested in a separate location from general education students.

Special education students must take the NJ ASK unless their IEP specifically exempts them. A student whose IEP exempts them from taking the NJ ASK must participate in the APA. Special education students may be tested using accommodations/modifications specified in the students' Individualized Education Programs (IEPs) that are approved by the Office of State Assessments. Students who have a disability and are eligible under Section 504 of the Rehabilitation Act of 1973 may be tested using accommodations/modifications specified in the student's 504 plan that are approved by the Office of State Assessments.

Large-print and Braille materials are provided to districts as required. Students completing a Braille version of the Mathematics section are instructed to bring a Braille ruler to the test session as well as a talking calculator. Students completing a large-print version of the test may use a ruler that is used during class instruction.

Students using the Braille test booklets are permitted to dictate their answers for multiple-choice questions to the examiner. Students taking the Braille test are also permitted to dictate their responses to the open-ended questions and all writing tasks. If dictation is used, the student is required to indicate all punctuation and must spell all key words.

Students using the large-print test booklets mark their answers for multiple-choice questions in the large-print version of the test booklet. Visually impaired students may use special equipment such as a typewriter or computer, if appropriate, for the open-ended questions and writing tasks. In 2007, the Braille versions differed from the standard versions for all grades and subjects. One or more operational items were omitted from the standard version to create the Braille version of each form. These items are noted in the student's copy of the test. A list is provided to the

examiners along with the supplemental instructions for administering the large-print and Braille versions of the test.

Accommodations for Limited English Proficient Students

NCLB prohibits exemptions from testing based on limited English proficient (LEP) status. However, LEP students were tested with one or more accommodations in the test administration procedures. Permitted accommodations include the following:

- additional time up to 150% of the administration times indicated
- translation of the test directions only into the student's native language (translations of passages, items, prompts, and tasks are NOT permitted)
- use of a bilingual dictionary

Students who received translated test directions were tested in a location separate from students tested with directions read in English only.

PART 4: SCORING

4.1 Multiple Choice Items

Before documents are scanned, a complete check of the scanning system is conducted. A mock set of answer documents are gridded to cover all response ranges, demographic data, blanks, double grids and other responses. Mock student records are created to verify that each gridding possibility is processed correctly by the scanning program. The output file that is created is thoroughly hand-checked against each answer document after each stage to ensure that the scanners are capturing all marks correctly. When the program output is confirmed to match the expected results, a formal sign-off process takes place.

The scoring keys are reviewed and approved prior to entry into the scoring system, and once entered, are verified. The multiple-choice scoring process entails multiple reviews for accuracy performed by independent staff on each key in every form.

4.2 Open Ended Items

Scoring of Open-Ended (OE) items involves having trained scorers read each student response. The student responses are assigned points by the scorers based on rules outlined in scoring rubrics. Before 2007, all OE items were scored independently by two scorers and final scores were obtained by averaging the two ratings together. A decision was made by NJDOE in summer 2006 to change the scoring practices of NJ ASK OE items from two scorers to one. Prior to the 2007 test administration, a special study was conducted to determine what, if any, impact changes to the scoring process would have on student scores. A summary of this study is presented in section 7.5 of this technical report. The results indicated that the impact would be

minimal; however, the results also suggested that when moving to one scorer, quality control (QC) procedures become even more important. Additional OE item quality control procedures were implemented during the 2007 administration. For example, requirements for qualifying and monitoring scorers were increased, to make sure students were being treated as fairly as possible. Also, for reading, math and science, ten percent of the student responses receive a second score for quality control purposes.

All writing composition items continue to receive two scores due to their complexity. The student responses are assigned points by the scorers based on rules outlined in scoring rubrics. For more information about the scoring rubrics, readers are referred to the Cycle I and II Score Interpretation Manual at the following website:

<http://www.nj.gov/njded/assessment/es/njask2005manual.pdf>.

Scorer Selection

The selection of scorers for the constructed response items is made from a large pool of candidates who meet stringent qualifications. Scorers must have, at a minimum, a four-year college degree, and must complete an individual interview. Preference is given to individuals with degrees and backgrounds related to language arts, mathematics and/or science, and experience in performance scoring. If appropriate, they are also asked to complete a grammar placement test and submit an original writing sample. Scoring supervisors are chosen based on subject area expertise, along with strong organizational abilities and communication skills. Scoring supervisors must demonstrate the ability to assist scoring directors in training, calibration and discussion sessions by successfully articulating the unique scoring criteria and their application.

Range Finding

Range-finding sessions are conducted using a range of photocopied student responses for each item. These responses are used to expand and refine existing anchor sets (selected examples of student work representing the score points), to be used in the training for operational scoring.

Scorer Training

Comprehensive training for scorers is provided via an online training system. This system incorporates scoring guides, fully annotated sample responses, practice exercises and qualifying sets. The training is user-driven and interactive and scorers are able to set their own pace.

The scoring guides present the rubrics with descriptions of each score level, and guidelines are provided on how to properly apply the scoring criteria. Annotated papers are chosen to clearly represent each designated score point. These student responses serve as the primary points of reference for scorers as they internalize the rubric during training. All scorers have access to this anchor set whenever they are scoring, and are directed to refer to it regularly.

Practice sets of student responses are used during training to help scorers become more experienced in applying the rubric. The use of these practice sets provides guidance to scorers in defining the line between score points and in applying the scoring criteria to a wider range of types of responses.

Qualification sets of student responses incorporate a range of student performance levels to confirm that the trainees can correctly assign the full range of scores. Candidates must demonstrate acceptable performance on these sets in order to qualify as a scorer.

Calibration sets are utilized throughout scoring to re-emphasize the score points, the lines between score points and any other scoring issues that may be identified by the scoring staff. During the scoring of Reading, calibration sets are administered to the scorers at the beginning of the shift and/or after lunch break for the first four days of scoring. For all content areas, calibration is a part of the quality management procedures when an issue is being addressed with an individual scorer or a group of scorers.

Scoring Procedures

Once trained, the scorers review and score responses using an electronic scoring system, which is accessible from multiple locations. The security protocols within the system are designed to ensure that the individual who received the training and is qualified to score is the individual who is scoring the responses. Scoring rate, reliability and validity statistics are monitored by the system and by supervisors to identify changes or trends in the scorer's performance. If a scoring anomaly is suspected, the problematic scorer can be locked from the system and all, or a portion of their work, may be reset to address a scoring quality issue.

The system assigns priority to student responses within the pool of available student responses based on a first-in and first-out system, and delivers to the scorer the next eligible response from the pool. Items requiring second reads are given priority over unscored responses, and the system prevents a response from receiving the first and second scores from the same scorer.

During the scoring of writing, if the first and second scores for a response are non-adjacent (e.g., one reader assigns a "5", and the second reader a "3"), the response will automatically be forwarded to a scoring supervisor, who will review and score the response to resolve the discrepancy.

Qualified scorers are authorized to assign valid score points or the code "No Response." Supervisory staff score responses sent to them for review, responses with non-adjacent scores and all other responses requiring condition codes (Off Topic, Not English, Wrong Format).

4.3 Quality Control Procedures in Data Preparation

All information gridded on the students' test booklets is automatically scanned and a series of edit checks are applied during and after the scanning process, prior to storage of the data in a

master database. Some student demographic data in the database may be modified through an online password-protected system accessible to specified individuals within the districts.

The master database is the origination of all data for files and reports for the testing administration. This includes all paper reporting, reporting via on line password-protected system, and files for the preparation of other State reporting.

Each time data is extracted from the master database for any of the reporting cycles or other files required by the DOE, the extracted data is put through a series of quality control checks to ensure its accuracy for that reporting cycle or file. Once the extracted data has been verified as correct and complete, the reporting cycle continues with the production of reports or files.

PART 5: STANDARD SETTING

The NJ ASK currently assesses two subject areas in grade 3, Language Arts Literacy and Mathematics, and three subjects in grade 4, Language Arts Literacy, Mathematics, and Science. The Language Arts Literacy tests consist primarily of open-ended items, writing tasks, and some multiple-choice items. The Mathematics tests are primarily multiple-choice items and some open-ended items. The grade 4 Science assessment is similar to the mathematics design - primarily multiple-choice with a few open-ended items. For each grade and subject, standard setting workshops were conducted shortly after the tests were administered for the first time operationally (i.e., in the base year). There were no standard setting workshops held after the March 2007 administration.

In all cases, the cut scores are used to distinguish performance among three levels: Partially Proficient, Proficient, and Advanced Proficient. Once raw score cuts were established on the base forms, item response theory (IRT) equating procedures have been used to maintain the cuts over time on new forms. See Part 7, Scaling and Equating, for more information about equating procedures. Following is a brief description of the standard setting procedures used by ETS to set standards on the NJ ASK tests.

5.1 Grade 4 Language Arts Literacy and Mathematics

The NJ ASK was introduced in 2003. Grade 4 NJ ASK Language Arts Literacy and Mathematics were intended to be a continuation of the grade 4 Elementary School Proficiency Assessment (ESPA). The base year for grade 4 Mathematics was 1999 and for Language Arts Literacy was 2001. To maintain program consistency, the content, number of items, and number of score points were all consistent with the previous ESPA forms. Cut scores were set in the base years for each subject, when the program was called the ESPA. The reader should contact the NJDOE for more information about the standard setting procedures used to set cut scores on the grade 4 Language Arts Literacy and Mathematics forms.

5.2 Grade 3 Language Arts Literacy and Mathematics

After the March 2004 administration, standard setting workshops were held in June for the grade 3 Language Arts Literacy (LAL) and Mathematics tests. ETS conducted the standard setting workshops in two phases. The following text comes from the executive summary of the standard setting report. For more information about the standard setting workshop, the full report is available from the NJDOE.

Overview

The grade 3 Language Arts Literacy and Mathematics standard setting was conducted in two phases. Phase 1 involved approximately 20 educators from across the state of New Jersey meeting for 2 ½ to 4 days and using a research-based standard setting method to recommend cut scores. Phase 2 immediately followed Phase 1 and involved 3 teachers from each of the two Phase 1 panels as well as 3 additional policymakers from the state. They reviewed the Phase 1 cut scores along with additional information about the percentage of students who would be classified in each level. This additional information included the percentage of students in all reporting categories (e.g., economically disadvantaged) who would reach Proficient and Advanced Proficient and the percentages currently reaching those levels in grade 4. They then provided their recommendations for cut scores, which were presented to the New Jersey State Board of Education for review and adoption. Following is a brief summary of the procedure and the results.

Summary of the Process

In May 2004, two groups of educators were invited to participate in a one-day workshop to develop the performance level descriptors for Proficient and Advanced Proficient in LAL and mathematics. These descriptors were used throughout the standard setting process.

Two different methods were used to set standards because the two subject areas had different test specifications. Because the LAL test was comprised primarily of open-ended items and writing prompts, a holistic method was chosen to determine cut scores. The Body of Work method required panelists to review entire student booklets, including responses to both open-ended and multiple-choice items, and determine whether the skills and knowledge demonstrated in the booklet best match the performance level descriptors for Partially Proficient, Proficient, or Advanced. The panelists were not told what the scores were for each booklet, but the standard setting facilitators used the information on judges' ratings in combination with the scores for each booklet to calculate a cut score for each level. Body of Work was conducted over two rounds. In the first round, rangefinding, panelists were given 30 booklets with scores ranging from 4 to 38 points out of 40. Based on the ratings of these 30 booklets, a second set of booklets were pulled for round 2, the pinpointing round. After the rangefinding round, the preliminary cut scores were calculated to be 20 points for Proficient and 30 points for Advanced Proficient. Another 22 booklets were selected to cover the range of 15 to 25 points for the Partially Proficient/Proficient cut score and 22 more booklets at 25 to 35 points for the Proficient/Advanced Proficient cut score. These 44 new booklets were used in the Pinpointing round to determine exactly where within the initial ranges the cut scores should fall.

For mathematics, which consisted primarily of multiple-choice items, an item mapping procedure called Item-Descriptor (ID) Matching was used. This is a variant of the Bookmark procedure that orders the operational items by difficulty as determined by the scale location of the items. Thus, the items that students performed best on appear first in an ordered test booklet and the items they performed worst on appear last. Panelists first go through the ordered test booklet and match the knowledge and skills required by the each item to the knowledge and skills listed in the performance level descriptors. That is, they ask themselves what one has to know and be able to do to answer an item correctly and then determine whether those knowledge and skills more closely match the descriptions of Partially Proficient, Proficient, and Advanced Proficient. Once they have matched each item to a performance level descriptor, they then determine the location of the cut score that best separates Partially Proficient performance from Proficient performance and Proficient from Advanced Proficient performance. ID Matching is conducted over three rounds with panelists receiving feedback about their ratings and having a chance to discuss their ratings with their peers between rounds.

In both procedures, panelists received “consequence” information about the percentage of students that took the test in March that would be categorized as Partially Proficient, Proficient, or Advanced Proficient. The percentages given to the panelists were based on the cut scores set after the first Pinpointing rating in Body of Work and after Round 2 in ID Matching. They then had the chance to discuss this information with their peers and make final adjustments to their ratings before the conclusion of Phase 1.

Summary of Results

Overall, panelists’ judgments about the cut scores converged from one round to the next, showing strong agreement by the end of the final round of Phase 1 and even more agreement in Phase 2. Table 5.2.1 shows the recommended cut scores at the end of Round 2 (the equivalent of the first Pinpointing rating in Body of Work) before the panelists saw the consequences data, at the end of Round 3 and then after Phase 2. The standard error of measurement (SEM) shows the degree of uncertainty in a student’s score on the test around the cut score, and the standard error of judgment (SEJ) is related to the variance in panelists’ judgments around the cut score. Overall, we find the SEJs decreasing across rounds, indicating converging opinions. Table 5.2.1 also shows that the Phase 2 panelists adopted the Phase 1 recommendations for three of the four cut scores and only modified the recommendation for the proficient cut score in mathematics by 2 points—within 1 SEM of the cut score recommended at the end of Phase 1.

TABLE 5.2.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Grade 3 LAL and Mathematics Standard-Setting Results from 2004
Recommended Cut Scores at the End of Phase 1 and Phase 2**

	Phase 1 - Round 2		Phase 1 - Round 3		Phase 2	
	Proficient	Advanced Proficient	Proficient	Advanced Proficient	Proficient	Advanced Proficient
LAL						
Cut Score	19.5	32.5	18	30.5	18	30.5
SEM	2.5	2.0	2.5	2.0	2.5	2.0
SEJ	0.30	0.40	0.14	0.32	0.09	0.00
Math						
Cut Score	14.5	32.0	15	27.5	17.0	27.5
SEM	2.5	2.0	2.5	2.0	2.5	2.0
SEJ	0.50	0.31	0.32	0.35	0.25	0.12

Table 5.2.2 shows the final cut scores that were brought to the State Board of Education for their review and approval. The four cut scores recommended by the Phase 2 panel were presented along with the consequences data showing the percentage of grade 3 students who would be categorized as Partially Proficient, Proficient, and Advanced Proficient. In addition to these tables, the State Board was also provided with consequence data for students in each reporting category, such as gender, race/ethnicity, and economic status. The State Board voted unanimously to adopt the recommended cut scores for the grade 3 NJ ASK.

TABLE 5.2.2

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Grade 3 LAL and Mathematics Standard-Setting Results from 2004
Recommended Cut Scores After Phase 2**

	Cut score for Proficient	Cut Score for Advanced Proficient	% Partially Proficient	% Proficient	% Advanced Proficient
LAL	18.0	30.5	21.6%	74.6%	3.8%
Math	17.0	27.5	23.8%	53.4%	22.8%

5.3 Grade 4 Science

The Science program became operational in 2005. As a result, a standard setting workshop was held after the March 2005 administration to determine the cut scores for Science. ETS conducted the standard setting workshop in two phases. The following text comes from the executive

summary of the standard setting report. For more information about the standard setting workshop, the full report is available from the NJDOE.

Overview

The Science standard setting was conducted in two phases. Phase 1 involved 18 educators from across the State of New Jersey meeting for 2 days and using a research-based standard setting method to recommend cut scores. Phase 2 immediately followed Phase 1 and involved 3 teachers from the Phase 1 panel, as well as 3 additional policymakers from the state. The Phase 2 panelists reviewed the Phase 1 cut scores along with additional information about the percentage of students who would be classified in each level. This additional information included the percentage of students in all reporting categories (e.g., economically disadvantaged) who would reach Proficient and Advanced Proficient and the percentages currently reaching those levels in grade 4 Language Arts Literacy and mathematics. They then provided their recommendations for cut scores, which was presented to the New Jersey State Board of Education for review and adoption on July 6, 2005. Following is a brief summary of the procedure and the results.

Summary of the Process

In April 2005, a group of educators was invited to participate in a one-day workshop to develop the performance level descriptors for Proficient and Advanced Proficient in Science. These descriptors were used throughout the standard setting process. Since the Science test is predominately multiple-choice, but also integrates information from open-ended items, ETS proposed an extended Angoff method (Hambleton & Plake, 1995). The Angoff method is the most thoroughly researched method used in setting standards. Although the use of the Angoff method with NAEP was subject to some criticism (National Academy of Education, 1993, p.xxiv), the method was subsequently defended and continues to be strongly supported by prominent psychometricians (c.f., Cizek, 1993; Kane, 1995; Mehrens, 1995; Loomis & Bourke, 2001). The modified Angoff method continues to be the most commonly used method of setting cut scores for tests that are predominantly multiple-choice.

For each MC item, the panel was instructed to read each question, consider the minimally Proficient (and minimally Advanced Proficient) student, and rate each item as to “How many of those 100 minimally Proficient (and How many of those 100 minimally Advanced Proficient) students would answer this item correctly?” They were instructed to give their ratings in intervals of five points (e.g., 25, 30, 35). Since, by chance, 25 out of 100 would likely answer a MC item correctly, the lowest reasonable rating for any MC item is 25. The maximum rating was limited to 95 as we do not expect perfection from any student.

For the open-ended item, panelists were asked to estimate the expected score for borderline students (i.e., the average score out of 3 possible points). The panel was instructed to read the OE question and rate each item by answering this question: “If 100 minimally Proficient students (and 100 minimally Advanced Proficient students) took this OE item, what would their average score be?” The panelists were not restricted to any particular increments because other values are possible. For example, a panelist might decide that 10 borderline students would probably skip the item and receive zero points, 10 would get 0.5 points, 30 would score 1.0 points, 40 would

score a 1.5, and ten would receive a score of 2.0. Such a combination would yield an average score of 1.15.

After round 2, panelists received “consequence” information about the percentage of students who took the test in March that would be categorized as Partially Proficient, Proficient, or Advanced Proficient. The percentages given to the panelists were based on the average cut scores as of Round 2. The panelists then had the chance to discuss this information with their peers and make final adjustments to their ratings before the conclusion of Phase 1.

Summary of Results

Overall, panelists’ judgments about the cut scores converged from one round to the next, showing little variance by the end of the final round of Phase 1. Table 5.3.1 shows the recommended cut scores at the end of Round 2 before the panelists saw the consequence data, at the end of Round 3, and then after Phase 2. The standard error of measurement (SEM) shows the degree of uncertainty in a student’s score on the test around the cut score, and the standard error of judgment (SEJ) is related to the variance in panelists’ judgments around the cut score. Overall, we find the SEJs decreasing across rounds, indicating converging opinions. One column, Phase 1 – adjusted, reflects the cut scores adjusted for outlier effects. When the highest and lowest cut scores were removed (a common procedure for a modified Angoff), the resulting cut scores showed no change for Proficient, and an increase of 0.5 points for Advanced Proficient. Table 5.3.1 also shows that the Phase 2 panelists adopted the Phase 1 recommendations for Proficient, and recommended the cut score for Advanced Proficient that matched both the Round 2 rating and the Round 3 rating after it had been adjusted for outliers.

TABLE 5.3.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Grade 4 Science Standard-Setting Results from 2005
Recommended Cut Scores at the End of Phase 1 and Phase 2**

	Phase 1 - Round 2		Phase 1 - Round 3		Phase 1 - Adjusted		Phase 2	
	Advanced Proficient	Proficient	Advanced Proficient	Proficient	Advanced Proficient	Proficient	Advanced Proficient	Proficient
Science								
Cut Score	19	30	19	29.5	19	30	19	30
SEM	3.0	2.5	3.0	2.5	3.0	2.5	3.0	2.5
SEJ	0.52	0.53	0.50	0.52	0.47	0.28	0.00	0.41

Table 5.3.2 shows the final cut scores that were brought to the State Board of Education for their review and approval. In addition to these tables, the State Board was also provided with consequence data for students in each reporting category, such as gender, race/ethnicity, and economic status. The State Board voted unanimously to adopt the recommended cut scores for the NJ ASK grade 4 Science test.

TABLE 5.3.2

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Grade 4 Science Standard-Setting Results from 2005
Recommended Cut Scores After Phase 2**

	Cut score for Proficient	Cut Score for Advanced Proficient	% Partially Proficient	% Proficient	% Advanced Proficient
Science	19.0	30.0	19.6%	58.4%	22.0%

PART 6: TEST STATISTICS

6.1 Classical Item Statistics

For each administration, classical item analyses are completed prior to item calibration, scaling and equating. These statistics are calculated again once all of the data are available. These analyses involve computing, for every item in each form, a set of statistics based on classical test theory. Each statistic is designed to provide some key information about the quality of each item from an empirical perspective. The statistics estimated for the NJ ASK are described below.

- **Classical item difficulty (“P-Value”):**
This statistic indicates the percent of examinees in the sample that answered the item correctly. Desired p-values generally fall within the range of 0.30 to 0.90.
- **Item discrimination (“r-biserial”)¹:**
This statistic is measured by the polyserial correlation between the item score and the test criterion score and describes the relationship between performance on the specific item and performance on the entire form. The higher the value, the better the task of separating the examinees. Items with negative correlations can indicate serious problems with the item content (e.g., multiple correct answers or unusually complex content), or can indicate that students have not been taught the content. For Language Arts Literacy, the

¹ The estimated polyserial correlation between scores on the item and on the criterion is computed by the formula:

$$r_{polyreg} = \frac{\beta_i \sigma_x}{\sqrt{\beta_i^2 \sigma_x^2 + 1}},$$

where the β_i are a series of parameters estimated by maximum likelihood from the item analysis data (Drasgow, 1988; Lewis & Thayer, 1996).

test criterion score was the number-correct score on the MC items, plus the weighted CR item score. For mathematics, the test criterion score was the number-correct score.

- The proportion of students choosing each response option:
These statistics indicate the percent of examinees that select each of the available answer options and the percent of examinees that omitted the item.
- Distractor analyses for MC items.
The GENASYS system (GENASYS is a proprietary ETS item analysis software program) provides graphical displays of the data for each option, which are reviewed.
- Percent of students omitting an item:
This statistic is useful for identifying problems with test features such as testing time and item/test layout. Typically, we would expect that if students have an adequate amount of testing time, 95% of students should attempt to answer each question. When a pattern of omit percentages exceeds 5% for a series of items at the end of a timed section, this may indicate that there was insufficient time for students to complete all items. Alternatively, if the omit percentage is greater than 5% for a single item, this could be an indication of an item/test layout problem. For example, students might accidentally skip an item that follows a lengthy stem.

In Tables 6.1.1 and 6.1.2, summary statistics are given that describe the difficulty and discrimination of the items comprising each cluster for grades 3 and 4, respectively. For dichotomously scored items, means and standard deviations of proportion-correct values (p-values) and r-biserials are given. For the open-ended items, the index of item difficulty was calculated by dividing students' average scores on an item by the maximum possible score on the item. Item discrimination for each open-ended item is the correlation between students' item score and their total score on the test section. For both the item-test correlation and the r-biserial correlation, students' total test scores were expressed in terms of the raw score metric.

TABLE 6.1.1

2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)

**Item Difficulty and Discrimination Summary Statistics
for Dichotomously Scored and Open-Ended Items
by Test Section and Cluster – Grade 3**

NJ ASK Test Section/Cluster	Dichotomous			Open-Ended		
	Item Difficulty		Item Discrimination	Item Difficulty		Item Discrimination
	Mean	S.D.	Mean	Mean	S.D.	Mean
Language Arts Literacy	0.77	0.10	0.56	0.42	0.03	0.74
Writing	--	--	--	0.44	0.02	0.78
Writing/Picture	--	--	--	0.46	--	0.79
Writing/Poem	--	--	--	0.43	--	0.77
Reading	0.77	0.10	0.56	0.40	0.02	0.70
Working with Text	0.77	0.12	0.55	--	--	--
Analyzing Text	0.76	0.08	0.57	0.40	0.02	0.70
Mathematics	0.71	0.15	0.52	0.53	0.08	0.71
Number Sense & Numerical Operations	0.72	0.15	0.53	--	--	--
Geometry & Measurement	0.75	0.08	0.47	0.44	--	0.68
Patterns & Algebra	0.72	0.18	0.53	0.56	--	0.72
Data Analysis, Probability & Discrete Math	0.63	0.21	0.51	0.59	--	0.73
Problem Solving	0.73	0.15	0.57	0.53	0.08	0.71

TABLE 6.1.2

2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)

**Item Difficulty and Discrimination Summary Statistics
for Dichotomously Scored and Open-Ended Items
by Test Section and Cluster – Grade 4**

NJ ASK Test Section/Cluster	Dichotomous			Open-Ended		
	Item Difficulty		Item Discrimination	Item Difficulty		Item Discrimination
	Mean	S.D.	Mean	Mean	S.D.	Mean
Language Arts Literacy	0.73	0.09	0.55	0.41	0.09	0.70
Writing	--	--	--	0.50	0.06	0.75
Writing/Picture	--	--	--	0.54	--	0.77
Writing/Poem	--	--	--	0.46	--	0.72
Reading	0.73	0.09	0.55	0.35	0.01	0.68
Working with Text	0.71	0.10	0.56	--	--	--
Analyzing Text	0.75	0.08	0.53	0.35	0.01	0.68
Mathematics	0.69	0.15	0.53	0.51	0.08	0.69
Number Sense & Numerical Operations	0.82	0.07	0.53	0.55	0.10	0.66
Geometry & Measurement	0.62	0.16	0.48	0.55	--	0.70
Patterns & Algebra	0.63	0.16	0.56	0.54	--	0.77
Data Analysis, Probability & Discrete Math	0.61	0.11	0.56	0.39	--	0.66
Problem Solving	0.71	0.09	0.58	0.51	0.08	0.69
Science	0.67	0.14	0.51	0.53	0.04	0.60
Life Science	0.64	0.14	0.51	0.52	--	0.50
Physical Science	0.70	0.16	0.49	0.56	--	0.69
Earth Science	0.69	0.12	0.52	0.49	--	0.62
Application	0.66	0.14	0.51	0.53	0.04	0.60

Frequency distributions of the March 2007 NJ ASK item p-values (difficulty values) and item discrimination indices are provided by content section and cluster for Language Arts Literacy, Mathematics, and Science in Tables 6.1.3, 6.1.4, 6.1.5, 6.1.6, and 6.1.7. The top section of each table shows the distribution of item difficulty values; the bottom section shows the distribution of r-biserial indices.

TABLE 6.1.3**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Frequency Distributions of Item Difficulty Values and Biserial
Discrimination Indices by Content Cluster****Language Arts Literacy – Grade 3**

Item Statistics	Working With Text	Analyzing Text	Total
ITEM DIFFICULTY: P-VALUES			
.800+	4	2	6
.700 - .799	3	1	4
.600 - .699	0	1	1
.500 - .599	1	0	1
<.500	0	0	0
MEAN P-VALUE	0.77	0.76	0.77
MEDIAN P-VALUE	0.80	0.77	0.80
ITEM DISCRIMINATION: BISERIAL CORRELATIONS			
.50+	6	4	10
.40 - .49	2	0	2
.27 - .39	0	0	0
MEAN POINT-BISERIAL	0.55	0.57	0.56
MEDIAN POINT-BISERIAL	0.56	0.57	0.56
TOTAL NUMBER OF ITEMS	8	4	12

TABLE 6.1.4

2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)

**Frequency Distributions of Item Difficulty Values and Biserial
Discrimination Indices by Content Cluster**

Language Arts Literacy – Grade 4

Item Statistics	Working With Text	Analyzing Text	Total
ITEM DIFFICULTY: P-VALUES			
.800 - .899	1	2	3
.700 - .799	3	1	4
.600 - .699	2	1	3
.500 - .599	1	0	1
<.500	0	0	0
MEAN P-VALUE	0.71	0.75	0.73
MEDIAN P-VALUE	0.70	0.77	0.72
ITEM DISCRIMINATION: BISERIAL CORRELATIONS			
.50+	6	3	9
.40 - .49	1	1	2
.30 - .39	0	0	0
MEAN POINT-BISERIAL	0.56	0.53	0.55
MEDIAN POINT-BISERIAL	0.58	0.54	0.56
TOTAL NUMBER OF ITEMS	7	4	11

TABLE 6.1.5

2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)

**Frequency Distributions of Item Difficulty Values
and Biserial Discrimination Indices by Content Cluster**

Mathematics – Grade 3

Item Statistics	Number Sense & Numerical Operations	Geometry & Measurement	Patterns & Algebra	Data Analysis, Probability & Discrete Math	Problem Solving	Total Test
ITEM DIFFICULTY: P-VALUES						
.900+	1	0	1	0	0	2
.800 - .899	3	1	1	1	1	6
.700 - .799	4	3	0	1	1	8
.600 - .699	2	1	1	2	0	6
.500 - .599	1	0	2	0	1	3
<.500	1	0	0	1	0	2
MEAN P-VALUE	0.72	0.75	0.72	0.63	0.73	0.71
MEDIAN P-VALUE	0.77	0.76	0.68	0.61	0.75	0.75
ITEM DISCRIMINATION: BISERIAL CORRELATIONS						
.50+	8	3	4	4	2	19
.40 - .49	4	1	1	1	1	7
.30 - .39	0	1	0	0	0	1
.20 - .29	0	0	0	0	0	0
MEAN POINT-BISERIAL	0.53	0.47	0.53	0.51	0.57	0.52
MEDIAN POINT-BISERIAL	0.52	0.50	0.53	0.52	0.54	0.52
Total Number of Items	12	5	5	5	3	27

TABLE 6.1.6

2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)

**Frequency Distributions of Item Difficulty Values
and Biserial Discrimination Indices by Content Cluster**

Mathematics – Grade 4

Item Statistics	Number Sense & Numerical Operations	Geometry & Measurement	Patterns & Algebra	Data Analysis, Probability & Discrete Math	Problem Solving	Total Test
ITEM DIFFICULTY: P-VALUES						
.900 - .999	2	0	0	0	1	2
.800 - .899	5	1	1	0	0	7
.700 - .799	4	1	1	1	4	7
.600 - .699	0	3	1	4	5	8
.500 - .599	0	0	3	1	1	4
<.500	0	2	1	1	0	4
MEAN P-VALUE	0.82	0.62	0.63	0.61	0.71	0.69
MEDIAN P-VALUE	0.81	0.65	0.59	0.65	0.68	0.70
ITEM DISCRIMINATION: BISERIAL CORRELATIONS						
.50 - .59	7	4	6	6	10	23
.40 - .49	4	2	1	1	1	8
.30 - .39	0	1	0	0	0	1
.20 - .29	0	0	0	0	0	0
MEAN POINT-BISERIAL	0.53	0.48	0.56	0.56	0.58	0.53
MEDIAN POINT-BISERIAL	0.53	0.51	0.53	0.59	0.60	0.53
Total Number of Items	11	7	7	7	11	32

TABLE 6.1.7**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Frequency Distributions of Item Difficulty Values and Biserial
Discrimination Indices by Content Cluster****Science – Grade 4**

Item Statistics	Life Science	Physical Science	Earth Science	Application	Total Test
ITEM DIFFICULTY: P-VALUES					
.900+	0	0	0	0	0
.800 - .899	1	4	2	4	7
.700 - .799	2	1	2	4	5
.600 - .699	6	1	3	9	10
.500 - .599	1	2	1	3	4
<.500	2	1	1	4	4
MEAN P-VALUE	0.64	0.70	0.69	0.66	0.67
MEDIAN P-VALUE	0.67	0.76	0.69	0.69	0.69
ITEM DISCRIMINATION: BISERIAL CORRELATIONS					
.50+	7	4	6	14	17
.40 - .49	4	4	2	7	10
.30 - .39	1	1	1	3	3
.20 - .29	0	0	0	0	0
MEAN POINT-BISERIAL	0.51	0.49	0.52	0.51	0.51
MEDIAN POINT-BISERIAL	0.52	0.49	0.55	0.53	0.53
Total Number of Items	12	9	9	24	30

6.2 Speededness

The NJ ASK is intended to provide sufficient time for all students to respond to almost all of the questions. The percentage of students omitting an item provides information about speededness, although it must be kept in mind that students can omit an item for reasons other than speededness (for example, choosing to not put effort into answering a constructed response item). Thus, if the percentage of omits is low, that implies that there is little speededness; if a percentage of omits is high, speededness, as well as other factors, can be the cause.

Tables 6.2.1 and 6.2.2 present data concerning the extent to which students omitted items. Table 6.2.1 shows that the percentage of grade 3 students omitting the Reading multiple-choice items was very small while the percentage of students omitting the Reading open-ended items varied from 1.0% to 1.2%. Table 6.2.1 also shows the percentage of grade 3 students omitting each of the last two Mathematics multiple-choice items in each part and all Mathematics open-ended items. The percentage of grade 3 students omitting the Mathematics multiple-choice items ranged from 0.3% to 2.0%. The percentage of grade 3 students omitting the Mathematics open-ended items ranged from 0.7% to 1.5%.

Table 6.2.2 shows that the percentage of grade 4 students omitting the Reading multiple-choice items was very small while the percentage of students omitting the Reading open-ended items varied from 0.7% to 3.8%. Table 6.2.2 also shows the percentage of grade 4 students omitting each of the last two Mathematics multiple-choice items in each part and all Mathematics open-ended items. The percentage of grade 4 students omitting the Mathematics multiple-choice items ranged from 0.3% to 2.4%. The percentage of grade 4 students omitting the Mathematics open-ended items ranged from 0.8% to 5.4%. Table 6.2.2 also shows the percentage of grade 4 students omitting each of the last two Science multiple-choice items in each part and all Science open-ended items. The percentage of grade 4 students omitting the Science multiple-choice items ranged from 0.3% to 0.8%. The percentage of grade 4 students omitting the Science open-ended items ranged from 0.5% to 0.7%.

TABLE 6.2.1

2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)

**Percentage of Students Omitting the
Last Items of Each Test Part – Grade 3**

Test Section	Multiple - Choice		Open - Ended	
	Item Number	Percentage Omitting	Item Number	Percentage Omitting
Reading				
<u>First Part</u>	Item 5	0.2%		
	Item 6	0.4%	Item 7	1.2%
<u>Second Part</u>	Item 5	0.4%		
	Item 6	0.7%	Item 7	1.0%
Mathematics				
<u>Day 1</u>				
<u>First Part</u>	Item 2	0.3%		
	Item 3	1.1%		
<u>Second Part</u>	Item 5	0.6%		
	Item 6	1.4%		
<u>Third Part</u>	Item 12	1.0%		
	Item 13	1.0%	Item 14	1.5%
<u>Fourth Part</u>	Item 20	1.1%		
	Item 21	2.0%	Item 22	1.5%
<u>Fifth Part</u>	Item 28	0.5%		
	Item 29	0.6%	Item 30	0.7%

TABLE 6.2.2

2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)

**Percentage of Students Omitting the
Last Items of Each Test Part – Grade 4**

Test Section	Multiple – Choice		Open - Ended	
	Item Number	Percentage Omitting	Item Number	Percentage Omitting
Reading				
<u>First Part</u>	Item 4	0.2%	Item 6	0.7%
	Item 5	0.2%	Item 7	3.8%
<u>Second Part</u>	Item 5	0.3%		
	Item 6	0.6%	Item 7	1.1%
Mathematics				
<u>Day 1</u>				
<u>First Part</u>	Item 3	0.3%		
	Item 4	0.8%		
<u>Second Part</u>	Item 7	0.4%		
	Item 8	0.5%		
<u>Third Part</u>	Item 19	1.9%		
	Item 20	2.4%	Item 21	2.0%
<u>Fourth Part</u>	Item 26	0.6%	Item 28	1.4%
	Item 27	0.8%	Item 29	5.4%
<u>Day 2</u>				
<u>Fifth Part</u>	Item 34	1.6%	Item 36	0.8%
	Item 35	0.4%	Item 37	1.6%

TABLE 6.2.2 (continued)

2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)

**Percentage of Students Omitting the
Last Items of Each Test Part – Grade 4**

Test Section	Multiple – Choice		Open - Ended	
	Item Number	Percentage Omitting	Item Number	Percentage Omitting
Science				
<u>Day 1</u>	Item 9	0.3%		
<u>First Part</u>	Item 10	0.3%	Item 11	0.5%
<u>Second Part</u>	Item 20	0.4%		
	Item 21	0.8%	Item22	
<u>Third Part</u>	Item 31	0.4%		0.7%
	Item 32	0.8%	Item 33	0.6%

6.3 Intercorrelations

The Pearson product-moment correlation between student scores on the Language Arts Literacy and Mathematics content areas for grade 3 was .69; this correlation for grade 4 was also .69. The correlation between student scores on the grade 4 Science and Language Arts Literacy content areas, and between Science and Mathematics content areas were .70 and .77, respectively. Tables 6.3.1 and 6.3.2 show the correlations between students' scores in the major content clusters and item types. Tables 6.3.3 and 6.3.4 show the correlations between students' scores on the content clusters. The scores used for all correlations were expressed in the raw score metric.

Note that correlations between a content area and cluster within that content area are partially a function of the proportion of the content area that is made up of items from the given cluster. All else being equal, clusters that make up a higher proportion of a content area score will tend to have higher cluster-area correlations. For example, the correlation between Mathematics Total and Mathematics Multiple-Choice in Table 6.3.2 is quite high at .96 because 28 Mathematics Multiple-Choice points are part of the Mathematics Total 43 points.

In addition, correlations are partially a function of the number of items in the measures being correlated; for a given pair of traits, increasing the number of items tends to increase correlations because of the increase in score reliability. Therefore, the number of items in the content areas and clusters being correlated must be considered when their correlations are evaluated.

TABLE 6.3.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Intercorrelations Among Major Content Clusters and Item Types – Grade 3**

Major Content Clusters and Item Types	Major Content Clusters and Item Types							
	Language Arts Literacy (LAL)					Mathematics (MAT)		
	LAL	R	R MC	R OE	W	MAT	M MC	M OE
LAL Language Arts Literacy (40)								
R Reading (20)	.93							
R MC Reading Multiple-Choice (12)	.87	.95						
R OE Reading Open-ended (8)	.78	.79	.56					
W Writing (20)	.84	.59	.51	.54				
MAT Mathematics (33)	.69	.66	.62	.55	.54			
M MC Mathematics Multiple-Choice (24)	.66	.65	.61	.52	.51	.95		
M OE Mathematics Open-ended (9)	.58	.55	.50	.48	.47	.87	.67	

Number in Parentheses is the number of points.
Language Arts Literacy N=100,877; Mathematics N=101,800.

TABLE 6.3.2

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Intercorrelations Among Major Content Clusters and Item Types – Grade 4**

Major Content Clusters and Item Types	Major Content Clusters and Item Types							
	Language Arts Literacy (LAL)					Mathematics (MAT)		
	LAL	R	R MC	R OE	W	MAT	M MC	M OE
LAL Language Arts Literacy (43)								
R Reading (23)	.94							
R MC Reading Multiple-Choice (11)	.84	.93						
R OE Reading Open-ended (12)	.82	.84	.57					
W Writing (20)	.83	.58	.47	.57				
MAT Mathematics (43)	.69	.68	.64	.56	.52			
M MC Mathematics Multiple-Choice (28)	.67	.66	.62	.53	.50	.96		
M OE Mathematics Open-ended (15)	.63	.61	.57	.52	.48	.92	.76	

Number in Parentheses is the number of points.
Language Arts Literacy N=100,617; Mathematics N=101,310.

TABLE 6.3.2 (Continued)

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Intercorrelations Among Major Content Clusters and Item Types – Grade 4**

Major Content Clusters and Item Types	Major Content Clusters and Item Types							
	Language Arts Literacy (LAL)					Science (SCI)		
	LAL	R	R MC	R OE	W	SCI	S MC	S OE
LAL Language Arts Literacy (43)								
R Reading (23)	.94							
R MC Reading Multiple-Choice (11)	.84	.93						
R OE Reading Open-ended (12)	.82	.84	.57					
W Writing (20)	.83	.58	.47	.57				
SCI Science (39)	.70	.71	.69	.55	.48			
S MC Science Multiple-Choice (30)	.68	.70	.68	.53	.46	.98		
S OE Science Open-ended (9)	.57	.57	.53	.46	.42	.80	.65	

Number in Parentheses is the number of points.
Language Arts Literacy N=100,617; Science N=101,266.

TABLE 6.3.2 (Continued)

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Intercorrelations Among Major Content Clusters and Item Types – Grade 4**

Major Content Clusters and Item Types	Major Content Clusters and Item Types					
	Mathematics (MAT)			Science (SCI)		
Major Content Clusters and Item Types	MAT	M MC	M OE	SCI	S MC	S OE
MAT Mathematics (43)						
M MC Mathematics Multiple-Choice (28)	.96					
M OE Mathematics Open-ended (15)	.92	.76				
SCI Science (39)	.77	.75	.69			
S MC Science Multiple-Choice (30)	.74	.72	.66	.98		
S OE Science Open-ended (9)	.66	.63	.60	.80	.65	

Number in Parentheses is the number of points.
Mathematics N=101,310; Science N=101,266.

TABLE 6.3.3

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Intercorrelations Among Content Areas and Clusters – Grade 3**

Test Section/Cluster	Test Section/Cluster												
	LAL Language Arts Literacy						MAT Mathematics						
Test Section/Cluster	LAL	L1	L2	L3	L4	L5	L6	MAT	M1	M2	M3	M4	M5
LAL Language Arts Literacy (40)													
L1 Reading (20)	.93												
L2 Writing (20)	.84	.59											
L3 Writing / Picture (10)	.75	.54	.88										
L4 Writing / Poem (10)	.74	.50	.90	.57									
L5 Working with Text (8)	.81	.89	.48	.45	.42								
L6 Analyzing Text (12)	.87	.92	.57	.52	.49	.64							
MAT Mathematics (33)	.69	.66	.54	.49	.46	.58	.62						
M1 Number Sense and Numerical Operations(9)	.59	.57	.46	.42	.40	.50	.53	.85					
M2 Geometry and Measurement (8)	.55	.53	.43	.40	.37	.45	.49	.80	.57				
M3 Data Analysis, Probability and Discrete Math (8)	.56	.55	.43	.40	.37	.48	.52	.83	.61	.54			
M4 Patterns and Algebra (8)	.57	.55	.44	.41	.38	.47	.52	.84	.63	.56	.58		
M5 Problem Solving (12)	.62	.59	.49	.45	.42	.51	.56	.91	.71	.74	.77	.79	

Number in Parentheses is the number of points.
Language Arts Literacy N=100,877; Mathematics N=101,800.

TABLE 6.3.4

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Intercorrelations Among Content Areas and Clusters – Grade 4**

Test Section/Cluster	Test Section/Cluster												
	LAL Language Arts Literacy						MAT Mathematics						
Test Section/Cluster	LAL	L1	L2	L3	L4	L5	L6	MAT	M1	M2	M3	M4	M5
LAL Language Arts Literacy (43)													
L1 Reading (23)	.94												
L2 Writing (20)	.83	.58											
L3 Writing / Picture (10)	.75	.55	.86										
L4 Writing / Poem (10)	.69	.45	.88	.50									
L5 Working with Text (7)	.79	.87	.45	.44	.35								
L6 Analyzing Text (16)	.89	.93	.57	.54	.46	.64							
MAT Mathematics (43)	.69	.68	.52	.50	.41	.61	.62						
M1 Number Sense and Numerical Operations(13)	.61	.59	.48	.46	.38	.53	.55	.87					
M2 Geometry and Measurement (10)	.58	.57	.43	.41	.34	.51	.52	.84	.65				
M3 Data Analysis, Probability and Discrete Math (10)	.60	.60	.44	.42	.35	.54	.54	.86	.65	.64			
M4 Patterns and Algebra (10)	.60	.59	.44	.42	.35	.54	.54	.88	.69	.66	.69		
M5 Problem Solving (23)	.66	.65	.50	.48	.39	.58	.60	.96	.89	.74	.85	.85	

Number in Parentheses is the number of points.
Language Arts Literacy N=100,617; Mathematics N=101,310.

TABLE 6.3.4 (Continued)

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Intercorrelations Among Content Areas and Clusters – Grade 4**

Test Section/Cluster	Test Section/Cluster											
	LAL Language Arts Literacy							SCI Science				
Test Section/Cluster	LAL	L1	L2	L3	L4	L5	L6	SCI	S1	S2	S3	S4
LAL Language Arts Literacy (43)												
L1 Reading (23)	.94											
L2 Writing (20)	.83	.58										
L3 Writing / Picture (10)	.75	.55	.86									
L4 Writing / Poem (10)	.69	.45	.88	.50								
L5 Working with Text (7)	.79	.87	.45	.44	.35							
L6 Analyzing Text (16)	.89	.93	.57	.54	.46	.64						
SCI Science (39)	.70	.71	.48	.47	.38	.66	.64					
S1 Life Science (15)	.63	.64	.43	.41	.33	.59	.57	.90				
S2 Physical Science (12)	.62	.63	.43	.41	.34	.57	.56	.86	.66			
S3 Earth Science (12)	.59	.60	.41	.40	.32	.56	.54	.87	.66	.64		
S4 Application (33)	.69	.71	.48	.46	.37	.65	.63	.99	.90	.86	.85	

Number in Parentheses is the number of points.
Language Arts Literacy N=100,617; Science N=101,266.

TABLE 6.3.4 (Continued)

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Intercorrelations Among Content Areas and Clusters – Grade 4**

Test Section/Cluster	Test Section/Cluster										
	MAT Mathematics						SCI Science				
Test Section/Cluster	MAT	M1	M2	M3	M4	M5	SCI	S1	S2	S3	S4
MAT Mathematics (43)											
M1 Number Sense and Numerical Operations(13)	.87										
M2 Geometry and Measurement (10)	.84	.65									
M3 Data Analysis, Probability and Discrete Math (10)	.86	.65	.64								
M4 Patterns and Algebra (10)	.88	.69	.66	.69							
M5 Problem Solving (23)	.96	.89	.74	.85	.85						
SCI Science (39)	.77	.65	.66	.68	.67	.73					
S1 Life Science (15)	.67	.57	.58	.60	.58	.64	.90				
S2 Physical Science (12)	.67	.57	.57	.60	.59	.64	.86	.66			
S3 Earth Science (12)	.68	.57	.59	.60	.59	.64	.87	.66	.64		
S4 Application (33)	.76	.65	.65	.68	.66	.73	.99	.90	.86	.85	

Number in Parentheses is the number of points.
Mathematics N=101,310; Science N=101,266.

6.4 Item Bias Statistics

Following the classical item analyses, Differential Item Functioning (DIF) studies were completed. One of the goals of test development is to assemble a set of items that provides an estimate of a student's ability that is as fair and accurate as possible for all groups within the population. DIF statistics are used to identify those items that identifiable groups of students (e.g. females, African Americans, Hispanics) with the same underlying level of ability have different probabilities of answering correctly. If the item is differentially more difficult for an identifiable subgroup, the item may be measuring something different from the intended construct. However, it is important to recognize that DIF flagged items might be related to actual differences in relevant knowledge or skill (item impact) or statistical Type I error. As a result, DIF statistics are used to identify potential sources of item bias. Subsequent review by content experts and bias/sensitivity committees determines the source and meaning of any differences that are seen.

ETS used two DIF detection methods: the Mantel-Haenszel and standardization approaches. As part of the Mantel-Haenszel procedure, the statistic described by Holland & Thayer (1986), known as MH D-DIF, was used. This statistic is expressed as the differences between the focal and reference group performance after conditioning on total test score. This statistic is reported on the ETS delta scale, which is a normalized transformation of item difficulty (proportion correct) with a mean of 12 and a standard deviation of 4. Negative MH D-DIF statistics favor the reference group and positive values favor the focal group. The classification logic used for flagging items is based on a combination of absolute differences and significance testing. Items that are not statistically significantly different based on the MH D-DIF ($p > 0.05$) are considered to have similar performance between the two studied groups; these items are considered to be functioning appropriately. For items where the statistical test indicates significant differences ($p < 0.05$), the effect size is used to determine the direction and severity of the DIF. For the Language Arts Literacy OE items, the Mantel-Haenszel procedure was executed where item categories are treated as integer scores and a chi-square test was carried out with one degree of freedom. The male and white groups are considered as reference groups and the female and other ethnic groups are categorized as focal groups.

Based on these DIF statistics, items are classified into one of three categories and assigned values of A, B or C (see Table 6.4.1). Category A contains negligible DIF, Category B items exhibit slight or moderate DIF, and Category C items have moderate to large values of DIF. Negative values imply that conditional on the matching variable, the focal group has a lower mean item score than the reference group. In contrast a positive value implies that, conditional on the matching variable, the reference group has lower mean item score than the focal group. For constructed-response items the MH D-DIF is not calculated, but analogous flagging rules based on the chi-square statistic are applied, resulting in classification into A, B, or C DIF categories.

TABLE 6.4.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
DIF Categories**

DIF Category	Definition
A (negligible)	MH D-DIF not significantly different from zero, or has an absolute value less than one.
B (slight to moderate)	MH D-DIF is significantly different from zero, and is either a) less than 1.5, or b) not significantly different from one.
C (moderate to large)	MH D-DIF is significantly different from one, and has an absolute value greater than 1.5.

Operational items flagged for negative C (C-) DIF are reviewed by an expert DIF review panel consisting of NJDOE staff responsible for the NJ ASK, and external educators identified by NJDOE during the item review meetings, to ensure that the items are free from any bias before being used to produce final test scores.

6.5 Summary Statistics

Means and standard deviations of students' raw scores on each content area are given in Tables 6.5.1 (grade 3) and 6.5.2 (grade 4) for the March 2007 test. These data are based on the total student populations with valid scores described in Part 1 and Appendix A. Table 6.5.1 shows that grade 3 students' mean raw scores were 21.2 of 40 points for Language Arts Literacy, and 21.3 of 33 points for Mathematics. The table also shows the standard deviations of the raw scores for grade 3 were 5.1 on Language Arts Literacy and 6.6 on Mathematics. Table 6.5.2 shows that grade 4 students' mean raw scores were 22.1 of 43 points for Language Arts Literacy, 26.3 of 43 points for Mathematics, and 24.7 of 39 points for Science. The table also shows the standard deviations of the raw scores for grade 4 were 5.6 on Language Arts Literacy, 9.1 on Mathematics, and 7.0 on Science. Raw score to scale score conversion tables by content area are included in Appendix C. Also, frequency distributions of the scale scores by content area are shown in Appendix C.

TABLE 6.5.1**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Means and Standard Deviations of Students'
Raw Scores by Test Section – Grade 3**

TEST SECTION	Number of Points	Raw Scores Mean	Standard Deviation	Number Tested
Language Arts Literacy	40	21.2	5.1	100,877
Mathematics	33	21.3	6.6	101,800

TABLE 6.5.2**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Means and Standard Deviations of Students'
Raw Scores by Test Section – Grade 4**

TEST SECTION	Number of Points	Raw Scores Mean	Standard Deviation	Number Tested
Language Arts Literacy	43	22.1	5.6	100,617
Mathematics	43	26.3	9.1	101,310
Science	39	24.7	7.0	101,266

Means and Standard Deviations of Students' Raw Scores

Tables 6.5.3 and 6.5.4 report the means and standard deviations for students' obtained numbers of raw score points by cluster on the March 2007 tests. Table 6.5.3 shows that in Language Arts Literacy, grade 3 students' mean percent correct was 53.0% overall with 61.9% in Reading and 44.1% in Writing. The mean raw score on the writing/speculate task in response to a picture was 4.5 points out of a possible 10 points and the mean raw score on the writing/analyze task in response to a poem was 4.3 points out of a possible 10 points. The mean percents correct in the two Reading clusters—Working with Text and Analyzing/Critiquing Text—were 77.1% and 51.7%.

With respect to the grade 3 students' percent correct scores on the Mathematics content clusters, the data in Table 6.5.3 indicate that the mean percent correct ranged from 60.5% in Data Analysis, Probability, and Discrete Math to 68.2% in Number Sense and Numerical Operations. The mathematics items are also categorized as Problem Solving and Total. The mean percent correct was 57.3% for Problem Solving and 64.4% for Total.

Table 6.5.4 shows that in Language Arts Literacy, grade 4 students' mean percent correct was 51.4% overall with 52.8% in Reading and 49.7% in Writing. The mean raw score on the writing/speculate task in response to a picture was 5.4 points out of a possible 10 points and the mean raw score on the writing/analyze task in response to a poem was 4.5 points out of a possible 10 points. The mean percents correct in the two Reading clusters—Working with Text and Analyzing/Critiquing Text—were 70.7% and 45.0%.

With respect to the grade 4 students' percent correct scores on the Mathematics content clusters, the data in Table 6.5.4 indicate that the mean percent correct ranged from 53.8% in Data Analysis, Probability, and Discrete Math to 68.8% in Number Sense and Numerical Operations. The mathematics items are also categorized as Problem Solving and Total. The mean percent correct was 59.7% for Problem Solving and 61.1% for Total.

With respect to the grade 4 students' percent correct scores on the Science content clusters, the data in Table 6.5.4 indicate that the mean percent correct ranged from 60.8% in Life Science to 66.6% in Physical Science. The Science items are also categorized as Application and Total. The mean percent correct was 61.6% for Application and 63.4% for Total.

Tables 6.5.5 and 6.5.6 show the means and standard deviations for the students' raw scores and percent correct scores on the dichotomously scored items by NJ ASK Content Area. Tables 6.5.7 and 6.5.8 provide means and standard deviations for students' raw scores and percent correct scores on the open-ended items by cluster.

TABLE 6.5.3

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Means and Standard Deviations of Students' Raw Scores
and Percent Correct by Content Area – Grade 3**

NJ ASK Content Area	Number of Items		Number of Possible Points	Raw Score		Percent Correct	
	Multiple- Choice	Open- Ended		Raw Scores Mean	Standard Deviation	Mean	Standard Deviation
Language Arts Literacy	12	4	40	21.2	5.1	53.0	12.7
Writing	0	2	20	8.8	2.3	44.1	11.3
Writing/Picture	0	1	10	4.5	1.2	45.4	12.2
Writing/Poem	0	1	10	4.3	1.3	42.8	13.4
Reading	12	2	20	12.4	3.4	61.9	17.1
Working with Text	8	0	8	6.2	1.7	77.1	21.6
Analyzing Text	4	2	12	6.2	2.0	51.7	17.0
Mathematics*	27	3	33	21.3	6.6	64.4	20.1
Number Sense and Numerical Operations*	12	0	9	6.1	2.0	68.2	22.4
Geometry and Measurement	5	1	8	5.1	1.9	63.3	23.6
Patterns and Algebra	5	1	8	5.2	2.0	65.2	25.1
Data Analysis, Probability, and Discrete Math	5	1	8	4.8	2.1	60.5	26.2
Problem Solving	3	3	12	6.9	3.3	57.3	27.2

* Six multiple-choice items in the Number Sense and Numerical Operations cluster and in the Mathematics total raw score are counted as one-half point.

TABLE 6.5.4**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Means and Standard Deviations of Students' Raw Scores
and Percent Correct by Content Area – Grade 4**

NJ ASK Content Area	Number of Items		Number of Possible Points	Raw Score		Percent Correct	
	Multiple- Choice	Open- Ended		Raw Scores Mean	Standard Deviation	Mean	Standard Deviation
Language Arts Literacy	11	5	43	22.1	5.6	51.4	12.9
Writing	0	2	20	9.9	2.4	49.7	11.9
Writing/Picture	0	1	10	5.4	1.3	54.0	13.3
Writing/Poem	0	1	10	4.5	1.4	45.4	14.3
Reading	11	3	23	12.1	3.8	52.8	16.6
Working with Text	7	0	7	5.0	1.8	70.7	25.4
Analyzing Text	4	3	16	7.2	2.4	45.0	15.3
Mathematics*	32	5	43	26.3	9.1	61.1	21.3
Number Sense and Numerical Operations*	11	2	13	8.9	2.8	68.8	21.9
Geometry and Measurement	7	1	10	5.9	2.4	59.4	24.2
Patterns and Algebra	7	1	10	6.0	2.7	60.0	26.7
Data Analysis, Probability, and Discrete Math	7	1	10	5.4	2.6	53.8	26.3
Problem Solving	11	4	23	13.7	5.5	59.7	24.1
Science	30	3	39	24.7	7.0	63.4	18.1
Life Science	12	1	15	9.1	3.0	60.8	20.1
Physical Science	9	1	12	8.0	2.4	66.6	20.1
Earth Science	9	1	12	7.6	2.6	63.7	21.5
Application	24	3	33	20.3	6.1	61.6	18.4

* Eight multiple-choice items in the Number Sense and Numerical Operations cluster and in the Mathematics total raw score are counted as one-half point.

TABLE 6.5.5**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Means and Standard Deviations of Students' Raw Scores
and Percent Correct on the Dichotomously Scored Items
by Content Area – Grade 3**

NJ ASK Content Area	Number of Points	Raw Scores		Percent Correct	
		Mean	Standard Deviation	Mean	Standard Deviation
Language Arts Literacy	12	9.2	2.5	76.6	20.9
Writing ^a	--	--	--	--	--
Writing/Picture	--	--	--	--	--
Writing/Poem	--	--	--	--	--
Reading	12	9.2	2.5	76.6	20.9
Working with Text	8	6.2	1.7	77.1	21.6
Analyzing Text	4	3.0	1.1	75.7	26.8
Mathematics*	24	16.6	4.5	69.0	18.6
Number Sense and Numerical Operations*	9	6.1	2.0	68.2	22.4
Geometry and Measurement	5	3.8	1.2	75.3	23.3
Patterns and Algebra	5	3.6	1.2	71.3	24.0
Data Analysis, Probability and Discrete Math	5	3.1	1.3	61.9	25.0
Problem Solving	3	2.2	0.9	72.7	30.3

* Six items in the Number Sense and Numerical Operations cluster and in the Mathematics total raw score are counted as one-half point.

a. There were no dichotomously scored writing items.

TABLE 6.5.6**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Means and Standard Deviations of Students' Raw Scores
and Percent Correct on the Dichotomously Scored Items
by Content Area –Grade 4**

NJ ASK Content Area	Number of Points	Raw Scores		Percent Correct	
		Mean	Standard Deviation	Mean	Standard Deviation
Language Arts Literacy	11	8.0	2.5	72.3	23.0
Writing ^a	--	--	--	--	--
Writing/Picture	--	--	--	--	--
Writing/Poem	--	--	--	--	--
Reading	11	8.0	2.5	72.3	23.0
Working with Text	7	5.0	1.8	70.7	25.4
Analyzing Text	4	3.0	1.1	75.1	26.4
Mathematics*	28	18.7	5.6	66.7	19.9
Number Sense and Numerical Operations*	7	5.7	1.4	81.7	20.2
Geometry and Measurement	7	4.3	1.6	61.9	23.3
Patterns and Algebra	7	4.4	1.8	62.8	25.1
Data Analysis, Probability and Discrete Math	7	4.2	1.9	60.3	27.0
Problem Solving	11	7.8	2.6	70.5	23.7
Science	30	20.1	5.6	66.8	18.6
Life Science	12	7.6	2.6	63.1	21.9
Physical Science	9	6.3	1.8	70.2	20.3
Earth Science	9	6.2	2.0	68.5	22.5
Application	24	15.6	4.6	65.2	19.1

* Eight items in the Number Sense and Numerical Operations cluster and in the Mathematics total raw score are counted as one-half point.

a. There were no dichotomously scored writing items.

TABLE 6.5.7**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Means and Standard Deviations of Students' Raw Scores
and Percent Correct on the Open-Ended Items by Content Areas and Clusters – Grade 3**

NJ ASK Content Area	Number		Raw Scores		Percent Correct	
	Items	Points	Mean	Standard Deviation	Mean	Standard Deviation
Language Arts Literacy	4	28	12.0	3.2	42.9	11.3
Writing	2	20	8.8	2.3	44.1	11.3
Writing/Picture	1	10	4.5	1.2	45.4	12.2
Writing/Poem	1	10	4.3	1.3	42.8	13.4
Reading	2	8	3.2	1.3	39.7	16.4
Working with Text	0	0	--	--	--	--
Analyzing Text	2	8	3.2	1.3	39.7	16.4
Mathematics	3	9	4.7	2.7	52.2	30.5
Number Sense, and Numerical Operations	0	0	--	--	--	--
Geometry and Measurement	1	3	1.3	1.1	43.4	38.2
Patterns and Algebra	1	3	1.7	1.2	55.1	39.5
Data Analysis Probability and Discrete Math	1	3	1.7	1.3	58.0	42.2
Problem Solving	3	9	4.7	2.7	52.2	30.5

TABLE 6.5.8**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Means and Standard Deviations of Students' Raw Scores
and Percent Correct on the Open-Ended Items by Content Areas and Clusters – Grade 4**

NJ ASK Content Area	Number		Raw Scores		Percent Correct	
	Items	Points	Mean	Standard Deviation	Mean	Standard Deviation
Language Arts Literacy	5	32	14.1	3.7	44.2	11.5
Writing	2	20	9.9	2.4	49.7	11.9
Writing/Picture	1	10	5.4	1.3	54.0	13.3
Writing/Poem	1	10	4.5	1.4	45.4	14.3
Reading	3	12	4.2	1.8	35.0	14.7
Working with Text	0	0	--	--	--	--
Analyzing Text	3	12	4.2	1.8	35.0	14.7
Mathematics	5	15	7.6	4.2	50.6	27.8
Number Sense, and Numerical Operations	2	6	3.2	1.9	53.7	31.4
Geometry and Measurement	1	3	1.6	1.2	53.6	40.0
Patterns and Algebra	1	3	1.6	1.3	53.5	42.6
Data Analysis Probability and Discrete Math	1	3	1.2	1.2	38.7	39.0
Problem Solving	4	12	6.0	3.4	49.9	28.5
Science	3	9	4.7	2.0	52.1	22.3
Life Science	1	3	1.5	0.9	51.5	28.8
Physical Science	1	3	1.7	1.0	55.7	32.0
Earth Science	1	3	1.5	1.0	49.2	31.7
Application	3	9	4.7	2.0	52.1	22.3

PART 7: SCALING AND EQUATING

When tests are administered on multiple occasions, there is a need to create multiple forms. A test form is a set of test questions that is built according to a set of content and statistical test specifications (Millman and Greene, 1989). It is difficult to create two forms that are identical in difficulty. Kolen and Brennan (1995) define equating as a statistical process used to adjust scores on test forms so scores on the forms can be used interchangeably. For example, the level of knowledge and skills needed to obtain a score of 200 on the 2007 grade 4 NJ ASK Mathematics form must be the same level of knowledge and skills needed to obtain a 200 on the 1999 grade 4 NJ ASK Mathematics form. To facilitate the correct interpretation of scores from multiple forms, test scores are reported as scale scores. Each form of a test has its own raw-to-scale conversion. The scale scores are intended to be comparable across forms within a grade and subject. NJ ASK scale scores are not comparable across subjects (e.g., Language Arts Literacy and Mathematics) or grades (e.g., 3 and 4).

7.1 Scaling

The total scores in the 2007 NJ ASK Language Arts Literacy and Mathematics sections are reported as scale scores with a range of 100 to 300. Please note that 100 and 300 are a theoretical floor and ceiling and may not actually be observed. The scale score of 200 is the cut point between Partially Proficient and Proficient students. The scale score of 250 is the cut point between Proficient and Advanced Proficient students. The score ranges are as follows:

Partially Proficient	100-199
Proficient	200-249
Advanced Proficient	250-300

The scores of students who are included in the Partially Proficient level are considered to be below the state minimum level of proficiency. These students may need additional instructional support, which could be in the form of individual or programmatic intervention. It is important that districts consider multiple measures with all students before making decisions about students' instructional placement.

Scale scores for the NJ ASK tests are linearly related to the raw score metric of the base year. Thus, to obtain scale scores for each test, a set of scaling parameters are applied to the raw score metrics in the base years. The base year is the year the cut scores were set on the form. The base year for the grade 4 Language Arts Literacy test is 2001. For grade 4 Mathematics, the base year is 1999. For grade 3 Language Arts Literacy and Mathematics, 2004 is the base year. And, for grade 4 Science, the base year is 2005. Table 7.1.1 shows the scaling parameters for each test.

TABLE 7.1.1**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Scaling Parameters for Base Forms**

Grade	Subject	Base Year	Points	Slope	Intercept
3	Language Arts Literacy	2004	0-40	4.00000	128.0000
	Mathematics	2004	0-33	4.76190	119.0477
4	Language Arts Literacy	2001	0-43	4.34783	106.5217
	Mathematics	1999	0-43	4.16667	104.1666
	Science	2005	0-39	4.54545	113.6365

7.2 Equating Language Arts Literacy

The equating design used in grade 3 and grade 4 Language Arts Literacy is the same. The base year for grade 3 is 2004. Scores on the 2007 NJ ASK grade 3 Language Arts Literacy form were equated back to scores on the 2004 NJ ASK grade 3 Language Arts Literacy base form via 2006 anchored Rasch difficulty parameters and using IRT true score equating procedures. The grade 3 base year Language Arts Literacy raw score scale ranged from 0-40.0. The base year raw cut score for Proficient was 18.0 (200) and the raw cut score for Advanced Proficient was 30.5 (250). These raw cut scores were derived from a standard-setting workshop in 2004.

Scores on the 2007 NJ ASK grade 4 Language Arts Literacy form were equated back to scores on the 2001 Language Arts Literacy base form via 2006 anchored Rasch difficulty parameters and using IRT true score equating procedures. The grade 4 base year Language Arts Literacy raw score scale ranged from 0-43.0. The base year raw cut score for Proficient was 21.5 (200) and the raw cut score for Advanced Proficient was 33.0 (250). These raw cut scores were derived from a standard-setting workshop in 2001.

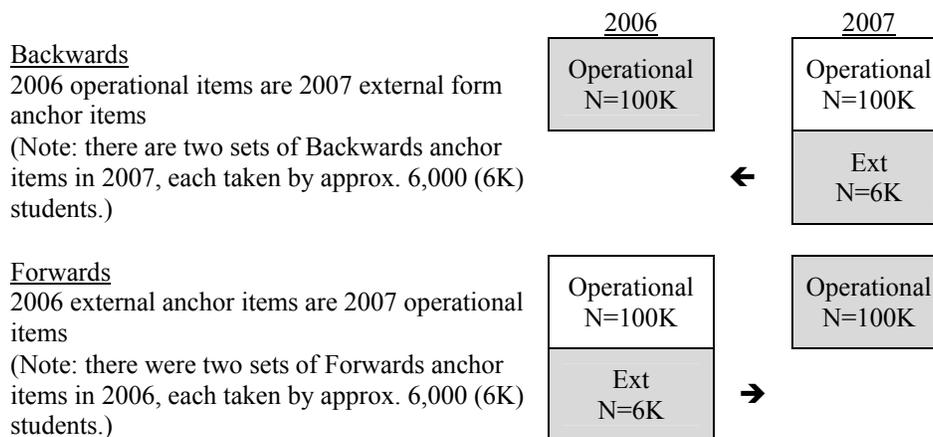
To perform equating, data must be collected. NJ ASK uses a Common-Item Nonequivalent Groups design. Common items are items that appear on both the reference (e.g., 2006) and new (e.g., 2007) forms. Common items are often also called linking and/or anchor items. The meaning of “Nonequivalent Groups” is that a different set of students took the reference and new forms, and no assumptions are made that the two groups are equal in ability. The groups could have the same ability, but the students taking the new form could also be more able or less able than the students taking the reference form.

The Language Arts Literacy equating design makes use of external anchor items (i.e., common items that do not count toward a student’s operational score). Language Arts Literacy uses an external anchor design that allows for two sets of anchor items to be used in the equating. The two designs have been called Backwards and Forwards. The Backwards equating anchor items were operational items on the old form (e.g., 2006) and are in external sets on the new form (e.g., 2007). The Forwards equating items were “pre-tested” as external sets on the old form (2006) and appear in the operational form on the new form (2007). In 2007, as recommended by the NJ

Technical Advisory Committee, the results of these two approaches were then combined to yield the final results.

Figure 7.2.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Language Arts Literacy Backwards and Forwards Equating Designs**



The final, Combined, equating approach makes use of the difficulty values from both Backwards and Forwards calibrations. The Backwards and Forwards difficulties are averaged. In addition, the step parameters are averaged. These item parameters are fixed and used to generate a Combined test characteristic curve (TCC). Through interpolation back to the base year a raw-score to scale-score conversion is obtained. The averaged b-values and step parameters will be used for equating in the following year.

For grade 3, performance on the equating anchor items indicates students in 2007 were about the same in ability as students in 2006, and the 2007 form was less difficult compared to the 2006 form. The recommended raw-score cut points 2007 for the grade 3 Language Arts Literacy test were 17 and 28 for proficient and advanced proficient categories, respectively. Details about the methods and results are described in the 2007 NJ ASK Grade 3 Language Arts Literacy Equating Report.

For grade 4, performance on the equating anchor items indicates students in 2007 were similar in ability compared to students in 2006, and the 2007 form was about the same difficulty as the 2006 form. The recommended raw-score cut points this year for Language Arts Literacy were 18 and 30 for proficient and advanced proficient categories respectively. Details about the methods and results are described in the 2007 NJ ASK Grade 4 Language Arts Literacy Equating Report.

Table 7.2.1 shows the Rasch difficulty parameters (“Measure”), and item fit statistics from WINSTEPS for the Combined equating solution for grade 3. Table 7.2.2 shows the fixed step parameters for the open-ended anchor items for grade 3. Table 7.2.3 shows the Rasch difficulty parameters, and item fit statistics from WINSTEPS for the Combined equating solution for grade

4. Table 7.2.4 shows the fixed step parameters for the open-ended anchor items for grade 4. The raw-to-scale score conversion tables for Language Arts Literacy for 2007 may be found in Appendix C.

To create Braille forms, a committee reviewed the 2007 Language Arts Literacy test items. Items that could not be translated into Braille were dropped from the Braille versions of the operational forms. In both grades 3 and 4 Language Arts Literacy the writing about pictures items (worth a maximum of 10 points) were dropped from the Braille forms. No other items were dropped. As a result, the Braille version of the grade 3 test was worth a maximum of 30 points (instead of 40) and the Braille version of the grade 4 test was worth a maximum of 33 points (instead of 43). Using the item parameters of the remaining items (in Tables 7.2.1 and 7.2.3), separate raw-to-scale score conversion tables were created for the Braille forms.

TABLE 7.2.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Language Arts Literacy Item Parameters – Grade 3**

Item No.	Measure	Anchor	Error	IN FIT		OUT FIT		Score Corr.	Displace
				MNSQ	ZSTD	MNSQ	ZSTD		
1	0.7984	Free	0.0019	0.86	-9.9	0.88	-9.9	0.76	0.02
2	0.3793	Anchor	0.0036	1.12	9.9	1.23	9.9	0.39	0.05
3	-0.4683	Anchor	0.0045	0.95	-9.3	0.91	-8.1	0.42	-0.04
4	-0.4783	Anchor	0.0045	0.92	-9.9	0.87	-9.9	0.42	-0.06
5	-0.1996	Anchor	0.0041	0.90	-9.9	0.85	-9.9	0.47	-0.08
6	-0.4293	Anchor	0.0044	0.88	-9.9	0.72	-9.9	0.50	-0.01
7	0.0632	Anchor	0.0038	0.96	-9.9	0.96	-6.2	0.48	-0.03
8	0.9942	Anchor	0.0024	0.90	-9.9	0.89	-9.9	0.65	0.07
9	0.9671	Free	0.0018	0.98	-3.1	1.03	4.8	0.74	-0.01
10	-0.1546	Anchor	0.0040	1.04	9.9	1.09	9.9	0.43	0.02
11	-0.9756	Anchor	0.0059	0.93	-8.5	1.22	9.9	0.29	-0.06
12	-0.5252	Anchor	0.0046	0.99	-2.7	0.97	-2.2	0.38	-0.05
13	-0.3395	Anchor	0.0043	0.98	-5.3	0.96	-3.8	0.43	-0.03
14	-0.1405	Anchor	0.0040	0.96	-9.9	0.92	-9.9	0.49	0.00
15	-0.4405	Anchor	0.0044	1.00	-0.7	1.03	2.7	0.41	0.00
16	1.1472	Anchor	0.0026	0.89	-9.9	0.89	-9.9	0.64	-0.03

TABLE 7.2.2

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Language Arts Literacy Fixed OE Item Step Parameters – Grade 3**

Item	Category	Step	Item	Category	Step
8	0	0.00	16	0	0.00
8	2	-2.84	16	2	-4.24
8	4	-1.62	16	4	-1.68
8	6	1.30	16	6	1.05
8	8	3.17	16	8	4.88

TABLE 7.2.3

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Language Arts Literacy Item Parameters – Grade 4**

Item No.	Measure	Anchor	Error	IN FIT		OUT FIT		Score Corr.	Displace
				MNSQ	ZSTD	MNSQ	ZSTD		
1	0.3177	Free	0.0017	0.93	-9.9	0.94	-9.9	0.75	0.01
2	-0.5020	Anchor	0.0044	0.89	-9.9	0.94	-6.3	0.42	-0.07
3	-0.5282	Anchor	0.0045	0.96	-8.1	0.92	-7.5	0.41	-0.02
4	0.0671	Anchor	0.0037	1.04	9.9	1.13	9.9	0.40	-0.07
5	0.0046	Anchor	0.0038	0.94	-9.9	0.90	-9.9	0.48	-0.06
6	-0.3132	Anchor	0.0041	1.00	0.9	1.14	9.9	0.40	-0.01
7	1.1636	Anchor	0.0026	0.96	-8.5	0.96	-8.4	0.67	0.04
8	1.0619	Anchor	0.0025	0.92	-9.9	0.91	-9.9	0.65	0.08
9	0.7302	Free	0.0017	1.17	9.9	1.26	9.9	0.69	0.00
10	-0.0907	Anchor	0.0038	0.96	-9.9	0.95	-8.0	0.47	-0.03
11	-0.4329	Anchor	0.0043	0.93	-9.9	0.91	-9.5	0.42	-0.06
12	-0.1057	Anchor	0.0039	0.99	-4.0	1.00	-0.5	0.42	-0.07
13	0.3489	Anchor	0.0036	1.10	9.9	1.18	9.9	0.38	-0.05
14	-0.3686	Anchor	0.0042	0.91	-9.9	0.87	-9.9	0.44	-0.07
15	0.0894	Anchor	0.0037	0.91	-9.9	0.89	-9.9	0.50	-0.07
16	1.2637	Anchor	0.0027	1.00	0.6	1.00	0.6	0.56	0.07

TABLE 7.2.4

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Language Arts Literacy Fixed OE Item Step Parameters – Grade 4**

Item	Category	Step		Item	Category	Step		Item	Category	Step
7	0	0.00		8	0	0.00		16	0	0.00
7	2	-3.88		8	2	-3.33		16	2	-4.24
7	4	-1.13		8	4	-1.57		16	4	-1.60
7	6	1.36		8	6	1.30		16	6	1.43
7	8	3.65		8	8	3.61		16	8	4.41

7.3 Equating Mathematics

The equating design used in grade 3 and grade 4 Mathematics is the same. The base year for grade 3 is 2004. Scores on the 2007 NJ ASK grade 3 Mathematics form were equated back to scores on the 2004 NJ ASK grade 3 Mathematics base form via 2006 anchored Rasch difficulty parameters and using IRT true score equating procedures. The grade 3 base year Mathematics raw score scale ranged from 0-33.0. The base year raw cut score for Proficient was 17.0 (200) and the raw cut score for Advanced Proficient was 27.5 (250). These raw cut scores were derived from a standard-setting workshop in 2004.

Scores on the 2007 NJ ASK grade 4 Mathematics form were equated back to scores on the 1999 Mathematics base form via 2006 anchored Rasch difficulty parameters and using IRT true score equating procedures. The grade 4 base year Mathematics raw score scale ranged from 0-43.0. The base year raw cut score for Proficient was 23.0 (200) and the raw cut score for Advanced Proficient was 35.0 (250). These raw cut scores were derived from a standard-setting workshop in 1999.

The data collection design for the NJ ASK Mathematics test is also a Common-Item Nonequivalent Groups design. The Mathematics test uses internal anchor items. Internal anchor items are common items that are embedded in the operational set of items (i.e., they count toward a student’s operational score).

For grade 3, equating was carried out using 11 anchor items from the 2006 form. Two anchor items were ½-point multiple-choice (MC) items, eight were 1-point MC items, and one was a three-point open-ended item (for a total of 12 points). All of the anchors were embedded in the new form. Sample size was 101,201 or approximately 99% of the total NJ grade 3 population with valid test scores. The 2007 students appear to be *similar in ability* compared to the 2006 students and the 2007 form was *more difficult* than the 2006 math form. The recommended raw-score (and scale-score) cut points for the 2007 Grade 3 Mathematics NJ ASK based on the equating results were 13.0 (200) and 25.5 (250) for Proficient and Advanced Proficient categories respectively.

For grade 4, equating was carried out using 12 anchor items from the 2006 form. The anchor set included two half-point and eight one-point multiple-choice items and two three-point open-ended items for a total of 15 points. All of the anchors were embedded in the new form. Sample size was 101,234 or approximately 99% of the total NJ grade 4 population with valid test scores. The 2007 students appear to be *slightly more able* than the 2006 students and the 2007 form was *more difficult* than the 2006 math form. The recommended raw-score (and scale-score) cut points for the 2007 Mathematics NJ ASK based on the equating results were 16.0 (200) for Proficient and 30.0 (250) for Advanced Proficient categories. Details about the methods and results are described in the 2007 NJ ASK Grade 4 Mathematics Equating Report.

Table 7.3.1 shows the Rasch difficulty parameters (“Measure”), and item fit statistics from WINSTEPS for the equating for grade 3. Table 7.3.2 shows the fixed step parameters for the open-ended items for grade 3. Table 7.3.3 shows the Rasch difficulty parameters, and item fit statistics from WINSTEPS for grade 4. Table 7.3.4 shows the fixed step parameters for the open-ended items for grade 4. The raw-to-scale score conversion tables for Mathematics for 2007 are presented in Appendix C. To create Braille forms a committee reviewed the 2007 Mathematics test items. Two items (15 and 22) were deleted from the grade 3 Braille form and one item (37) was deleted from the grade 4 Braille form. In addition, two items (8 and 15) were dropped from the grade 3 large print form. Therefore, separate raw-to-scale score conversion tables were created for the grades 3 & 4 Braille Mathematics forms and the grade 3 Mathematics large print form in 2007.

TABLE 7.3.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Mathematics Item Parameters – Grade 3**

Item No.	Measure	Anchor	Error	IN FIT		OUT FIT		Score	
				MNSQ	ZSTD	MNSQ	ZSTD	Corr.	Displace
1	-0.6575	Free	0.0082	0.92	-9.9	0.85	-9.9	0.35	-0.01
2	-1.6495	Free	0.0115	0.94	-7.2	0.79	-9.9	0.29	0.00
3	-0.2486	Anchor	0.0074	0.90	-9.9	0.85	-9.9	0.39	-0.03
4	-0.4479	Anchor	0.0077	0.90	-9.9	0.83	-9.9	0.40	-0.01
5	-0.4595	Free	0.0078	0.95	-9.9	0.92	-9.9	0.31	-0.01
6	-0.7079	Free	0.0083	0.92	-9.9	0.84	-9.9	0.36	-0.01
7	0.6673	Free	0.0036	1.05	9.9	1.08	9.9	0.40	0.00
8	0.1089	Free	0.0041	0.96	-8.8	0.92	-9.6	0.42	-0.01
9	0.2317	Free	0.0040	0.84	-9.9	0.75	-9.9	0.52	-0.01
10	-0.2373	Free	0.0049	0.95	-8.3	0.94	-4.7	0.37	-0.01
11	0.4368	Free	0.0037	1.01	4.0	1.08	9.9	0.41	-0.01
12	0.5954	Anchor	0.0036	1.05	9.9	1.06	9.9	0.41	0.03
13	1.3624	Free	0.0038	1.11	9.9	1.37	9.9	0.33	0.00
14	0.7447	Anchor	0.0018	1.04	8.6	1.04	7.0	0.65	0.00
15	0.5733	Free	0.0036	1.04	9.9	1.05	9.1	0.40	0.00
16	1.2061	Free	0.0036	0.98	-7.7	1.07	9.9	0.45	0.00
17	-0.0348	Free	0.0044	0.88	-9.9	0.72	-9.9	0.47	0.00

Item No.	Measure	Anchor	Error	IN FIT		OUT FIT		Score	
				MNSQ	ZSTD	MNSQ	ZSTD	Corr.	Displace
18	0.1923	Anchor	0.0040	1.11	9.9	1.08	9.9	0.40	0.10
19	0.6542	Free	0.0036	0.99	-2.8	0.98	-3.7	0.44	0.00
20	0.6353	Anchor	0.0036	0.98	-8.3	0.97	-6.2	0.47	0.05
21	-0.4368	Anchor	0.0055	1.03	4.2	0.97	-2.2	0.36	0.07
22	1.0414	Free	0.0019	1.13	9.9	1.15	9.9	0.62	-0.01
23	0.1803	Anchor	0.0040	1.17	9.9	1.36	9.9	0.24	-0.04
24	-0.2295	Anchor	0.0049	1.04	5.9	1.03	2.3	0.35	0.03
25	-0.7278	Free	0.0067	0.97	-3.2	0.98	-0.8	0.28	0.00
26	0.7985	Free	0.0035	1.02	9.0	1.05	9.9	0.42	0.00
27	0.5875	Free	0.0036	1.03	9.9	1.06	9.9	0.41	-0.01
28	0.5540	Anchor	0.0036	1.05	9.9	1.05	8.9	0.41	0.05
29	0.1397	Anchor	0.0041	1.05	9.9	1.10	9.9	0.36	0.01
30	0.7161	Free	0.0018	1.12	9.9	1.16	9.9	0.63	0.00

TABLE 7.3.2

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Mathematics OE Item Step Parameters – Grade 3**

Item	Category	Step
14	0	0.00
14	2	-0.64
14	4	0.77
14	6	-0.13

Item	Category	Step
22	0	0.00
22	2	-0.26
22	4	-0.43
22	6	0.69

Item	Category	Step
30	0	0.00
30	2	0.01
30	4	0.48
30	6	-0.49

TABLE 7.3.3

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Mathematics Item Parameters – Grade 4**

Item No.	Measure	Anchor	Error	IN FIT		OUT FIT		Score	
				MNSQ	ZSTD	MNSQ	ZSTD	Corr.	Displace
1	-0.4699	Free	0.0081	0.94	-9.9	0.86	-9.9	0.35	0.00
2	-0.2279	Free	0.0076	0.90	-9.9	0.84	-9.9	0.41	0.00
3	-2.0430	Free	0.0144	0.96	-3.3	0.81	-9.9	0.23	0.00
4	-0.5547	Anchor	0.0083	0.87	-9.9	0.78	-9.9	0.38	-0.08
5	-0.4917	Free	0.0081	0.90	-9.9	0.80	-9.9	0.41	0.00
6	-1.1104	Free	0.0099	0.94	-9.2	0.83	-9.9	0.31	0.00
7	-1.0982	Anchor	0.0099	0.91	-9.9	0.83	-9.9	0.30	-0.04
8	0.0736	Free	0.0071	0.90	-9.9	0.87	-9.9	0.42	0.00
9	0.5584	Anchor	0.0038	1.03	7.6	1.03	4.9	0.42	0.02
10	0.9550	Free	0.0035	1.04	9.9	1.09	9.9	0.40	0.00
11	0.9937	Free	0.0035	0.88	-9.9	0.86	-9.9	0.53	0.00
12	0.9477	Anchor	0.0035	1.02	6.9	1.03	7.2	0.44	0.12
13	0.5889	Free	0.0037	0.94	-9.9	0.90	-9.9	0.48	0.00
14	0.3227	Free	0.0041	1.03	6.8	1.04	4.7	0.39	0.00
15	-0.3244	Free	0.0056	0.90	-9.9	0.76	-9.9	0.39	0.00
16	0.0820	Free	0.0045	1.12	9.9	1.27	9.9	0.28	0.00
17	1.2077	Free	0.0036	1.03	9.9	1.10	9.9	0.41	0.00
18	0.7151	Anchor	0.0036	0.96	-9.9	0.93	-9.9	0.47	0.01
19	0.4156	Anchor	0.0039	0.89	-9.9	0.80	-9.9	0.50	0.00
20	1.2882	Free	0.0036	1.10	9.9	1.28	9.9	0.34	0.00
21	1.1191	Anchor	0.0020	0.92	-9.9	0.92	-9.9	0.67	-0.02
22	0.6161	Free	0.0037	1.12	9.9	1.18	9.9	0.34	0.00
23	0.6075	Anchor	0.0037	0.94	-9.9	0.90	-9.9	0.50	0.03
24	0.5697	Free	0.0037	1.08	9.9	1.12	9.9	0.37	0.00
25	0.3150	Free	0.0041	0.94	-9.9	0.87	-9.9	0.46	0.00
26	1.4058	Free	0.0037	1.15	9.9	1.33	9.9	0.30	0.00
27	-0.2075	Free	0.0053	0.96	-5.8	0.97	-1.8	0.36	0.00
28	0.9594	Anchor	0.0018	0.95	-9.9	0.91	-9.9	0.69	0.01
29	1.2659	Free	0.0019	1.20	9.9	1.25	9.9	0.60	0.00
30	0.6188	Anchor	0.0037	1.03	9.9	1.05	8.3	0.42	0.02
31	0.9667	Free	0.0035	1.00	-0.8	1.02	5.0	0.44	0.00
32	0.7914	Free	0.0036	0.95	-9.9	0.92	-9.9	0.48	0.00
33	0.8435	Anchor	0.0036	0.95	-9.9	0.94	-9.9	0.47	-0.05
34	0.3273	Free	0.0040	0.94	-9.9	0.89	-9.9	0.46	0.00
35	0.2399	Anchor	0.0042	0.91	-9.9	0.84	-9.9	0.45	-0.02
36	0.8190	Free	0.0018	1.58	9.9	2.10	9.9	0.53	0.00
37	0.9688	Free	0.0018	1.10	9.9	1.13	9.9	0.64	0.00

TABLE 7.3.4

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Mathematics OE Item Step Parameters – Grade 4**

Item	Category	Step
21	0	0.00
21	2	-1.06
21	4	0.15
21	6	0.91

Item	Category	Step
28	0	0.00
28	2	0.45
28	4	-0.50
28	6	0.06

Item	Category	Step
29	0	0.00
29	2	-0.15
29	4	-0.10
29	6	0.25

Item	Category	Step
36	0	0.00
36	2	0.18
36	4	0.71
36	6	-0.88

Item	Category	Step
37	0	0.00
37	2	-0.03
37	4	-0.34
37	6	0.37

7.4 Equating Science

The NJ ASK grade 4 Science test became operational in 2005. Standard setting workshops were held after the administration. See Part 5, Standard-Setting, for more information about the scaling of the 2005 NJ ASK grade 4 Science section.

The base year for grade 4 Science is 2005. Scores on the 2007 NJ ASK grade 4 Science form were equated back to scores on the 2005 base form via the 2006 anchored Rasch difficulty parameters and using IRT true score equating procedures. The grade 4 Science base year raw score scale ranged from 0-39.0. The base year raw cut score for Proficient was 19.0 (200) and the raw cut score for Advanced Proficient was 30.0 (250). These raw cut scores were determined at the standard-setting workshop in 2005.

The data collection design for the NJ ASK Science test is also a Common-Item Nonequivalent Groups design. The Science test uses internal anchor items. Internal anchor items are common items that are embedded in the operational set of items (i.e., they count toward a student's operational score).

Equating was carried out using 11 anchor items from the 2006 form. Ten anchor items were multiple-choice and one was open-ended (for a total of 13 points). All of the anchors were embedded in the new form. Sample size was 101,199 or approximately 99% of the total New Jersey grade 4 population. The 2007 students appear to be *more able* than the 2006 students and the 2007 form was *more difficult* than the 2006 science form. The recommended raw-score (and scale-score) cut points for the 2007 NJ ASK Science test based on the equating results were 18.0 (200) and 28.0 (250) for Proficient and Advanced Proficient categories, respectively. Details about the methods and results are described in the 2007 NJ ASK Grade 4 Science Equating Report.

Table 7.4.1 shows the Rasch difficulty parameters (“Measure”), and item fit statistics from WINSTEPS for the equating for grade 4 Science. Table 7.4.2 shows the fixed step parameters for the open-ended items for grade 4 Science. The raw-to-scale score conversion tables for Science 2007 are presented in Appendix C.

To create a Braille form, a committee reviewed the Science test items. Items that could not be translated into Braille were dropped from the Braille version of the operational form. In 2007, two items could not be translated into Braille (i.e., items 7, and 22). One item was multiple-choice and one was open-ended (worth a maximum of 4 points). As a result, the Science Braille version of the grade 4 test was worth a maximum of 35 points (instead of 39). Using the item parameters of the remaining items (in Table 7.4.1), a separate raw-to-scale score conversion table was created for the Braille form.

TABLE 7.4.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Science Item Parameters – Grade 4**

Item No.	Measure	Anchor	Error	IN FIT		OUT FIT		Score Corr.	Displace
				MNSQ	ZSTD	MNSQ	ZSTD		
1	0.4326	Free	0.0035	1.05	9.9	1.05	9.9	0.38	-0.01
2	0.0259	Free	0.0037	0.90	-9.9	0.83	-9.9	0.49	-0.01
3	0.0145	Free	0.0037	1.09	9.9	1.16	9.9	0.31	-0.01
4	0.6026	Anchor	0.0035	0.99	-3.7	1.02	4.4	0.42	0.04
5	-0.6440	Anchor	0.0050	0.86	-9.9	0.70	-9.9	0.39	-0.04
6	-0.5601	Anchor	0.0047	1.02	4.1	1.06	5.8	0.30	0.01
7	-0.3852	Anchor	0.0043	0.97	-5.9	0.94	-7.3	0.42	0.05
8	-0.0263	Anchor	0.0037	1.19	9.9	1.24	9.9	0.28	0.07
9	0.1221	Free	0.0036	0.91	-9.9	0.86	-9.9	0.49	-0.01
10	0.2567	Anchor	0.0035	1.09	9.9	1.12	9.9	0.35	0.05
11	0.4956	Free	0.0022	1.20	9.9	1.20	9.9	0.48	0.00
12	0.2973	Free	0.0035	1.07	9.9	1.09	9.9	0.35	0.00
13	0.7374	Free	0.0035	1.11	9.9	1.18	9.9	0.32	-0.01
14	-0.2218	Free	0.0040	0.95	-9.9	0.88	-9.9	0.41	-0.01
15	-0.1456	Free	0.0039	0.91	-9.9	0.81	-9.9	0.47	-0.01
16	0.9516	Anchor	0.0037	1.06	9.9	1.12	9.9	0.36	-0.02
17	-0.3533	Free	0.0042	0.91	-9.9	0.81	-9.9	0.43	-0.01
18	0.3094	Free	0.0035	0.96	-9.9	0.94	-9.9	0.45	-0.01
19	0.1013	Free	0.0036	0.92	-9.9	0.88	-9.9	0.47	-0.01
20	-0.1779	Free	0.0039	0.97	-7.7	0.97	-3.8	0.39	-0.01
21	0.0341	Free	0.0037	0.96	-9.9	0.91	-9.9	0.44	-0.01
22	0.5643	Anchor	0.0021	1.03	6.8	1.03	6.4	0.59	0.02
23	-0.4472	Anchor	0.0044	1.00	0.8	0.97	-3.6	0.35	0.02
24	0.6490	Free	0.0035	1.09	9.9	1.15	9.9	0.33	-0.01
25	-0.2518	Anchor	0.0041	0.97	-8.1	0.87	-9.9	0.45	0.05
26	-0.3969	Free	0.0043	1.07	9.9	1.24	9.9	0.26	0.00

Item No.	Measure	Anchor	Error	IN FIT		OUT FIT		Score	
				MNSQ	ZSTD	MNSQ	ZSTD	Corr.	Displace
27	0.1943	Free	0.0036	1.01	3.0	1.00	0.3	0.40	0.00
28	-0.1559	Free	0.0039	0.89	-9.9	0.80	-9.9	0.48	-0.01
29	0.0890	Free	0.0036	0.94	-9.9	0.88	-9.9	0.46	-0.01
30	-0.3758	Free	0.0043	0.94	-9.9	0.90	-9.9	0.40	-0.01
31	0.0029	Anchor	0.0037	1.06	9.9	1.09	9.9	0.34	-0.01
32	0.0067	Free	0.0037	0.94	-9.9	0.88	-9.9	0.45	-0.01
33	0.3293	Free	0.0021	0.92	-9.9	0.92	-9.9	0.65	-0.01

TABLE 7.4.2

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Science OE Item Step Parameters – Grade 4**

Item	Category	Step
11	0	0.00
11	2	-1.62
11	4	-0.36
11	6	1.97

Item	Category	Step
22	0	0.00
22	2	-0.75
22	4	-1.00
22	6	1.76

Item	Category	Step
33	0	0.00
33	2	-1.42
33	4	0.19
33	6	1.23

7.5 Rescore Equating Study

A decision was made in summer 2006 to change the way NJ ASK open-ended (OE) items were to be scored. Previously, each OE item received scores from two raters. Beginning in 2007, scores for OE items would be based on one rater only and, therefore, ½-point values would no longer assigned. (Note: The writing composition items for LAL, scored by summing the scores of two raters, did not change.)

Historically, OE items have been used as anchor items in NJ ASK equating. Adjusting the scoring process of an anchor OE item by using one rater instead of two raters is equivalent to using a “new” item. The following questions needed to be addressed in order to understand the impact of this change to the anchor item process:

1. How will changing scoring from two raters to one rater affect student scores and score distributions?
2. If “replacement” OE anchor items are used, are the equating results obtained substantially different?
3. How will this shift in OE scoring procedures affect the cut scores and the percentage of students in each performance level?

To answer these questions, ETS conducted a Rescore Equating Study using data from the 2005 and 2006 administrations— data were reanalyzed using results for one rater only. The 2006 equatings were replicated and the resulting conversions were compared to the operational

conversions used to report scores in 2006 (AKA the ‘old’ conversions). Following is a brief summary of the answers to the questions of interest.

1. How will changing scoring from two raters to one rater affect student scores and score distributions?

The results demonstrated that the state mean and overall score distributions did not change when students were scored using only rater one (R1) or rater two (R2) scores for OE items. In fact, the correlation between scores in all cases was greater than .99 and the reliability varied little. In addition, the cumulative frequency curves overlapped completely.

However, scores for individual students did change depending on whether the R1 or the R2 scores were used. Some scores increased while others decreased, although for science, less than 2% changed by more than +/- 2 raw score points in 2006. In order to protect the interests of students it was recommended that automatic rescoring for OE items be extended to students within two raw score points of the Proficient cut score.

2. If “replacement” OE anchor items are used, are the equating results obtained substantially different?

For Math and Science cut scores did not change, although, the raw-to-scale score conversions changed at some points (due to rounding). The LAL scales were impacted to a greater extent than were Math and Science primarily because there are more OE item points on the LAL tests and the Reading OE items (which consist of 5 categories: 0, 1, 2, 3, 4) generally have lower exact agreement rates. As a result of switching to one rater only, the LAL theta scales showed greater spread.

It was concluded that the 2006 OE items could be used as anchor items for the 2007 administration if they were recalibrated and put on scale prior to the May 2007 equating process. We did this by treating the OE items as if they were field-test items. We fixed all operational MC items to their 2006 operational values. This is the procedure that was followed and worked in the NJ ASK Rescore Equating Study.

It was also necessary to recalibrate and update the conversion tables for all “breach” forms. A breach form is administered when there is a need for a student to take a back-up or a make-up form. Breach forms for NJ ASK are usually a previously administered test that already has an existing conversion table.

3. How will this shift in OE scoring procedures affect the cut scores and the percentage of students in each performance level?

Students can no longer obtain ½-point scores in Science and LAL, and the cut scores may fall on the ½-point values, consequently it was recommended that the rounding rule be re-evaluated. After re-evaluation, the old rounding rules remained in place for Mathematics (because ½ point values were still possible). Science and LAL rounding rules were adjusted slightly and are described in detail in the equating reports.

The Rescore Equating Study was designed to evaluate the feasibility of moving from the scoring of OE items with two raters to a single rater. It was determined that if item parameters were recalibrated and treated as “new” items, the impact was minimal. However, it was also clear that when moving to one rater quality control (QC) procedures become critically important. Additional OE item quality control procedures were implemented during the 2007 administration including a rater year effect study (Fitzpatrick, Ercikan, Yen, & Ferrara, 1989).

Part 8: VALIDITY

Content and Curricular Validity

The New Jersey Department of Education is developing a comprehensive set of assessments that measure student achievement of the Core Curriculum Content Standards. The validity of the NJ ASK scores is based on the alignment of the NJ ASK assessments to the Core Curriculum Content Standards and the knowledge and skills expected of third- and fourth-grade students.

The Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999, p. 11-12) notes the following possible sources of validity evidence:

- Evidence based on test content
- Evidence based on internal structure of the test
- Evidence based on relations to other variables
- Evidence based on consequences of testing

For an assessment like NJ ASK, which is intended to measure students’ performance in relation to the Core Curriculum Content Standards, content validity evidence is primary. Content validity is the most relevant and important source of evidence. The section of this technical report on “Test Development,” presents validity evidence based on test content. A description of the test specification development is followed by the procedures for test item development. Details about item writing as well as task, prompt, and passage selection are included. The last section delineates the review work of the New Jersey Assessment Content Committees. Additionally, an external committee is assisting the New Jersey Department of Education by reviewing the assessments to determine how well they measure the knowledge and skills stated in the standards, and by comparing the New Jersey standards with those in other states and countries.

PART 9: TEST RELIABILITY

9.1 Classical Reliability Estimates of the Test Scores

Tables 9.1.1 and 9.1.2 summarize reliability estimates for the NJ ASK grades 3 and 4 content areas and clusters. The reliability coefficients given in these tables are based on Cronbach's coefficient alpha measure of internal consistency. Cronbach's alpha is used on tests containing items that can be scored along a range of values. The standard errors of measurement (SEMs) for the major content areas - Language Arts Literacy and Mathematics - are expressed in terms of the raw score metric and the scale score metric. The NJ ASK scale scores range from 100 to 300.

Reliabilities and SEMs for the dichotomously scored items in each cluster are reported in Tables 9.1.3 and 9.1.4.

When evaluating these results, it is important to recall that reliability is partially a function of test length. Therefore, the reliability of a content area is likely to be greater than the reliability of a cluster simply because the content area has more items. Similarly, clusters with more items are likely to be more reliable than clusters with fewer items. The data provided in Tables 9.1.1, 9.1.2, 9.1.3 and 9.1.4 reflect the expected positive relationship between test length and reliability.

The SEMs given in Tables 9.1.1, 9.1.2, 9.1.3 and 9.1.4 are useful when interpreting students' scores. Measurement error occurs in every test. A student's true score is a hypothetical average score that the student would obtain if a test were repeatedly administered to the student without the effects of instruction, practice, or fatigue. Mehrens and Lehmann (1991) suggest this use of the SEM:

The standard error of measurement is often used for what is called band interpretation. Band interpretation helps convey the idea of imprecision of measurement.... If we assume that the errors are random, an individual's observed scores will be normally distributed about his true score over repeated testing. Thus, one can say that a person's observed score will lie between ± 1 SE of his true score approximately 68 percent of the time, or ± 2 SE of his true score about 95 percent of the time (p. 252).

TABLE 9.1.1**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Reliability Estimates and Standard Errors of Measurement (SEM)
for Content Areas and Clusters – Grade 3**

NJ ASK Test Section	Number of Points	Reliability	Raw Score SEM	Scale Score SEM
Language Arts Literacy	40	0.83	2.07	9.24
Reading	20	0.79	1.55	.
Writing	20	0.73	1.19	.
Working with Text	8	0.64	1.04	.
Analyzing Text	12	0.68	1.15	.
Mathematics	33	0.86	2.53	10.72
Number Sense and Numerical Operations	9	0.70	1.11	.
Geometry and Measurement	8	0.45	1.40	.
Patterns and Algebra	8	0.51	1.40	.
Data analysis, Probability and Discrete Math	8	0.49	1.50	.
Problem Solving	12	0.56	2.17	.

TABLE 9.1.2**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Reliability Estimates and Standard Errors of Measurement (SEM)
for Content Areas and Clusters – Grade 4**

NJ ASK Test Section	Number of Points	Reliability	Raw Score SEM	Scale Score SEM
Language Arts Literacy	43	0.83	2.30	10.00
Reading	23	0.81	1.68	.
Writing	20	0.67	1.38	.
Working with Text	7	0.65	1.05	.
Analyzing Text	16	0.71	1.31	.
Mathematics	43	0.89	2.98	10.72
Number Sense and Numerical Operations	13	0.69	1.58	.
Geometry and Measurement	10	0.57	1.59	.
Patterns and Algebra	10	0.60	1.69	.
Data analysis, Probability and Discrete Math	10	0.65	1.55	.
Problem Solving	23	0.82	2.34	.
Science	39	0.86	2.66	12.27
Life Science	15	0.69	1.69	.
Physical Science	12	0.62	1.50	.
Earth Science	12	0.66	1.51	.
Application	33	0.83	2.48	.

TABLE 9.1.3**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Reliability Estimates and Standard Errors of Measurement (SEM)
for Dichotomously Scored Items Within Content Clusters – Grade 3**

NJ ASK Content Area	Number of Points	Reliability	Raw Score SEM
Language Arts Literacy	12	0.74	1.28
Reading	12	0.74	1.28
Writing*	--	--	--
Writing/Picture	--	--	--
Writing/Poem	--	--	--
Working with Text	8	0.64	1.04
Analyzing Text	4	0.50	0.76
Mathematics	24	0.82	1.92
Number Sense and Numerical Operations	9	0.70	1.11
Geometry and Measurement	5	0.42	0.89
Patterns and Algebra	5	0.47	0.87
Data analysis, Probability and Discrete Math	5	0.45	0.93
Problem Solving	3	0.44	0.68

TABLE 9.1.4**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)****Reliability Estimates and Standard Errors of Measurement (SEM)
for Dichotomously Scored Items Within Content Clusters – Grade 4**

NJ ASK Content Area	Number of Points	Reliability	Raw Score SEM
Language Arts Literacy	11	0.74	1.30
Reading	11	0.74	1.30
Writing*	--	--	--
Writing/Picture	--	--	--
Writing/Poem	--	--	--
Working with Text	7	0.65	1.05
Analyzing Text	4	0.46	0.78

TABLE 9.1.4 (Continued)
2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)

**Reliability Estimates and Standard Errors of Measurement (SEM)
for Dichotomously Scored Items Within Content Clusters – Grade 4**

NJ ASK Content Area	Number of Points	Reliability	Raw Score SEM
Mathematics	28	0.86	2.07
Number Sense and Numerical Operations	7	0.71	0.76
Geometry and Measurement	7	0.51	1.14
Patterns and Algebra	7	0.55	1.18
Data analysis, Probability and Discrete Math	7	0.64	1.13
Problem Solving	11	0.74	1.32
Science	30	0.83	2.29
Life Science	12	0.68	1.48
Physical Science	9	0.56	1.22
Earth Science	9	0.62	1.24
Application	24	0.80	2.07

* There were no dichotomously scored writing items.

9.2 Reliability of Performance Classifications

Decision accuracy provides an estimate of how reliably scores from a test form classify students into performance categories. It is determined by comparing the observed score distribution for a form to a hypothetical true score distribution. The observed score distribution (also called single-form score distribution) is the actual distribution of scores for all test takers on a test form. The true score distribution is hypothetical because true scores cannot be known, although, they can be estimated. A true score is the average of the observed scores for a student obtained over an infinite number of repeated administrations of the same form.

The methodology used for estimating the reliability of classification and decision accuracy is described in Livingston and Lewis (1995) and is implemented using the ETS-proprietary computer program RELCLASS-COMP (Version 4.12). RELCLASS-COMP generates a contingency table that shows the proportion of exact agreement between the two distributions. In Tables 9.2.1 and 9.2.2, the cells showing exact agreement are shaded. The sum of the shaded, diagonal cells represents the estimated proportion correctly classified.

Table 9.2.1: For grade 3 Language Arts Literacy, the estimated proportion correctly classified overall was 0.87. When the decisions were collapsed to below proficient versus proficient and above, the estimated proportion correctly classified was 0.93. For Mathematics, the estimated

proportion correctly classified overall was 0.83. When the decisions were collapsed to below proficient versus proficient and above, the estimated proportion correctly classified was 0.94.

Table 9.2.2: For grade 4 Language Arts Literacy, the estimated proportion correctly classified overall was 0.85. When the decisions were collapsed to below proficient versus proficient and above, the estimated proportion correctly classified was 0.92. For Mathematics, the estimated proportion correctly classified overall was 0.85. When the decisions were collapsed to below proficient versus proficient and above, the estimated proportion correctly classified was 0.94. For Science, the estimated proportion correctly classified overall was 0.81. When the decisions were collapsed to below proficient versus proficient and above, the estimated proportion correctly classified was 0.92.

TABLE 9.2.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Reliability of Classification and Decision Accuracy – Grade 3**

Decision Accuracy: Language Arts Literacy

		<i>Observed Score</i>			
		Advanced Proficient (28.0-40.0)	Proficient (17.0-27.0)	Partially Proficient (0-16.0)	Observed Total
<i>True Score</i>	Placement Score				
	Advanced Proficient (28.0-40.0)	0.00	0.07	0.01	0.08
	Proficient (17.0-27.0)	0.00	0.73	0.03	0.75
	Partially Proficient (0-16.0)	0.00	0.03	0.14	0.17
Expected Total		0.00	0.83	0.17	

Estimated Proportion Correctly Classified: Total = 0.87, Proficient & Above = 0.93

Decision Accuracy: Mathematics

		<i>Observed Score</i>			
		Advanced Proficient (25.5-33.0)	Proficient (13.0-25.0)	Partially Proficient (0-12.5)	Observed Total
<i>True Score</i>	Placement Score				
	Advanced Proficient (25.5-33.0)	0.26	0.06	0.00	0.32
	Proficient (13.0-25.0)	0.05	0.48	0.02	0.55
	Partially Proficient (0-12.5)	0.00	0.04	0.09	0.13
Expected Total		0.31	0.58	0.11	

Estimated Proportion Correctly Classified: Total = 0.83, Proficient & Above = 0.94

TABLE 9.2.2

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Reliability of Classification and Decision Accuracy – Grade 4**

Decision Accuracy Language Arts Literacy

		<i>Observed Score</i>			
		Advanced Proficient (30.0-43.0)	Proficient (18.0-29.0)	Partially Proficient (0-17.0)	Observed Total
Placement Score					
<i>True Score</i>	Advanced Proficient (30.0-43.0)	0.00	0.06	0.00	0.07
	Proficient (18.0-29.0)	0.01	0.69	0.04	0.74
	Partially Proficient (0-17.0)	0.00	0.04	0.16	0.19
	Expected Total	0.02	0.79	0.20	

Estimated Proportion Correctly Classified: Total = 0.85, Proficient & Above = 0.92

Decision Accuracy Mathematics

		<i>Observed Score</i>			
		Advanced Proficient (30.0-43.0)	Proficient (16.0-29.5)	Partially Proficient (0-15.5)	Observed Total
Placement Score					
<i>True Score</i>	Advanced Proficient (30.0-43.0)	0.36	0.05	0.00	0.41
	Proficient (16.0-29.5)	0.04	0.37	0.02	0.44
	Partially Proficient (0-15.5)	0.00	0.04	0.12	0.15
	Expected Total	0.40	0.46	0.14	

Estimated Proportion Correctly Classified: Total = 0.85, Proficient & Above = 0.94

TABLE 9.2.2 (Continued)

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Reliability of Classification and Decision Accuracy – Grade 4**

		Decision Accuracy Science			
		<i>Observed Score</i>			
	Placement Score	Advanced Proficient (28.0-39.0)	Proficient (18.0-27.0)	Partially Proficient (0-17.0)	Observed Total
<i>True Score</i>	Advanced Proficient (28.0-39.0)	0.33	0.07	0.01	0.41
	Proficient (18.0-27.0)	0.05	0.35	0.03	0.42
	Partially Proficient (0-17.0)	0.00	0.04	0.13	0.17
	Expected Total	0.38	0.46	0.16	

Estimated Proportion Correctly Classified: Total = 0.81, Proficient & Above = 0.92

9.3 Conditional Estimate of Error at Each Cut-Score

When reviewing a cut score, it is important to keep in mind that there is measurement error surrounding that cut score. Measurement error occurs because no instrument measures a student’s level of knowledge and skills precisely. Think of the student who knows the correct answer to an item, but makes a careless arithmetic error or accidentally marks the wrong response. Or think of a student who really does not know the correct answer but who fills in the correct answer purely by chance. These situations require us to calculate a standard error of measurement for each score. For example, let’s say a student scores a 200 and the standard error of measurement for the score is about 10 scale score points. We can be 95% confident that the student’s ability puts him in the range of scoring a 200 plus or minus two standard errors of measurement: that is between 180–220.

The WINSTEPS program calculates the standard error of the measure (SEM) at each score point. Unlike the classical standard error of measurement, the value of the SEM using Item Response Theory varies with ability level. The equation for standard error of estimation is given by

$$SE(\hat{\theta}) = \frac{1}{\sqrt{I(\theta)}} \tag{9.3.1}$$

where $I(\theta)$ is the information function for a test at θ . For the Rasch model using unweighted raw scores, the information provided by a test at θ is the sum of the item information functions at θ (Hambleton, Swaminathan, and Rogers, 1991). Table 9.3.1 shows conditional estimates of error at each cut score for each subject in grades 3 and 4.

TABLE 9.3.1**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Conditional Estimate of Error at Each Cut-Score**

Grade	Subject	Proficiency Level	Raw Score Cut	Theta Cut	Theta SE	Scale Score Cut	Estimated SE in Scale Score Points
3	LAL	Proficient	17.0	-0.0922	0.228	200	10
		Advanced Proficient	28.0	1.4528	0.297	250	7
	Math	Proficient	13.0	0.2344	0.203	200	13
		Advanced Proficient	25.5	1.1314	0.206	250	9
Grade	Subject	Proficiency Level	Raw Score Cut	Theta Cut	Theta SE	Scale Score Cut	Estimated SE in Scale Score Points
4	LAL	Proficient	18.0	-0.0543	0.220	200	10
		Advanced Proficient	30.0	1.3720	0.262	250	7
	Math	Proficient	16.0	0.4714	0.172	200	12
		Advanced Proficient	30.0	1.1695	0.163	250	10
	Science	Proficient	18.0	0.0162	0.171	200	14
		Advanced Proficient	28.0	0.6475	0.193	250	11

9.4 Rater Reliability

Beginning in 2007, scores for OE items were based on one rater only while writing composition items continued to be scored by summing the scores of two raters. To assess and maintain the quality of ratings of OE items, rater effects were analyzed using recognized methodology. This methodology involved rescoring a random sample of papers from 2006 and examining the degree of agreement between the original scores based on two raters and the rescores from a single rater. For each subject and grade level, a random sample of 1,000 (approximately 10%) papers from 2006 were obtained and rescored. Tables 9.4.1 and 9.4.2 show the percentages of writing tasks and the open-ended items scored with exact agreement, adjacent agreement, and resolution needed.

The Writing cluster within Language Arts Literacy consists of two writing activities: a writing/speculate task in response to a picture and a writing/analyze task related to a poem. For these writing tasks, the rubrics used by the raters had score points that ranged from 0 to 5. If two raters assigned scores to a student's writing task that were not exactly the same or adjacent, a third "expert" rater also read and assigned a score to the student's response. Of more than 200,000 writing task responses in grade 3 in March 2007, 64.3% received exactly the same scores by the raters and 34.3% received scores that were adjacent. Thus, about 98.5% of the task responses required only two raters. The remaining 1.5% received scores on the writing tasks that differed by more than one point and therefore required a third rater (see Table 9.4.1). For grade

4 Language Arts Literacy writing tasks in March 2007, 59.5% received exactly the same scores by the raters and 37.9% received scores that were adjacent. Thus, a total of 97.4% of the task responses required only two raters. The remaining 2.6% received scores on the writing tasks that differed by more than one point and therefore required a third rater (see Table 9.4.2).

The Reading cluster and the Mathematics content areas include open-ended items. For the Reading open-ended items, the rubric used by the raters had score points that ranged from 0 to 4. For the Mathematics items, the rubric ranged from 0 to 3 points. Table 9.4.1 shows that for grade 3 Reading open-ended items, exact agreement was obtained 61.4% of the time. Resolution by a third rater was needed for 2.2% of the responses. For grade 3 Mathematics, exact agreement was obtained 86.9% of the time and resolution was needed for 1.4% of the task responses. Table 9.4.2 shows that for grade 4 Reading open-ended items, exact agreement was obtained 62.1% of the time. Resolution by a third rater was needed for 1.9% of the responses. For grade 4 Mathematics, exact agreement was obtained 89.7% of the time and resolution was needed for 1.1% of the responses. Finally, for grade 4 Science, exact agreement was obtained 85.7% of the time and resolution was needed for 0.3% of the open-ended responses.

TABLE 9.4.1

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Consistency Between Raters Scoring Writing Tasks and Open-Ended Items - Grade 3**

Writing Tasks and Open-Ended Items	Percent Raters In Exact Agreement	Percent Raters In Adjacent Agreement	Percent Resolution Needed
Language Arts Literacy	62.8	35.3	1.8
Writing Total	64.3	34.3	1.5
Writing/Picture	63.9	34.6	1.5
Writing/Poem	64.6	33.9	1.5
Reading Total	61.4	36.4	2.2
Open-Ended Item 1	58.3	38.4	3.3
Open-Ended Item 2	64.6	34.4	1.0
Mathematics	86.9	11.8	1.4
Open-Ended Item 1	86.3	12.3	1.4
Open-Ended Item 2	80.6	17.8	1.5
Open-Ended Item 3	93.6	5.2	1.2

TABLE 9.4.2

**2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Consistency Between Raters Scoring Writing Tasks and Open-Ended Items – Grade 4**

Writing Tasks and Open-Ended Items	Percent Raters In Exact Agreement	Percent Raters In Adjacent Agreement	Percent Resolution Needed
Language Arts Literacy	61.1	36.8	2.2
Writing Total	59.5	37.9	2.6
Writing/Picture	60.1	37.6	2.2
Writing/Poem	58.8	38.2	3.0
Reading Total	62.1	36.0	1.9
Open-Ended Item 1	63.4	34.7	1.9
Open-Ended Item 2	61.8	36.2	2.0
Open-Ended Item 3	61.2	37.1	1.7
Mathematics	89.7	9.2	1.1
Open-Ended Item 1	86.2	13.6	0.3
Open-Ended Item 2	90.3	9.0	0.7
Open-Ended Item 3	89.3	8.7	2.0
Open-Ended Item 4	93.9	4.9	1.2
Open-Ended Item 5	88.5	9.8	1.6
Science	85.7	14.1	0.3
Open-Ended Item 1	77.3	22.3	0.5
Open-Ended Item 2	95.8	4.1	0.1
Open-Ended Item 3	84.0	15.8	0.2

Part 10: REPORTING

Scores are reported in two cycles, Cycle I and Cycle II. Cycle I data are considered preliminary. Schools and districts are encouraged to review student information to make sure it is correct and accurate before Cycle I reports are released. Schools have the opportunity to make corrections to student information before Cycle II reports are published. In addition, to minimize the risk of misclassification, Cycle I open ended items are automatically rescored for students who received a raw score within two points of a proficiency level cutoff score. Rescoring is also done at the request of districts. When the rescoring of a student's responses produces a higher raw score, the student's scale score is adjusted to reflect this change. Cycle II reports, which contain the rescored results are considered final.

10.1 Cycle I Reports

The Cycle I reports include the following: Student Sticker, Individual Student Report, All Sections Roster, Student Roster, Summary of School Performance, Summary of District Performance, Summary of School Cluster Performance, and Summary of District Cluster Performance. Each Cycle I report is briefly described below.

Student Sticker

The Student Sticker is produced alphabetically, and one sticker for each student within the school is provided. It is a peel-off label designed to be easily attached to the student's permanent record.

The scale scores in Language Arts Literacy, Mathematics and (for Grade 4 students) Science are provided. Designations of the proficiency levels are printed next to the Language Arts Literacy, Mathematics and Science scale scores. Voids, where applicable, are noted.

Individual Student Report

The Individual Student Report (ISR) is a two-sided report, produced in alphabetical sequence for students within the school. Two copies of this report are produced for every student tested, one for the student's permanent folder after the results are analyzed, and the other for the student's parent/guardian to be shared in a manner determined by the local district.

The scale scores in Language Arts Literacy, Mathematics and Science are provided on the front of the ISR (Figure 10.1.1), along with explanatory text about scale scores and proficiency levels. Cluster data is provided on the back of the ISR (Figure 10.2.1), along with explanatory text about cluster scores.

The Just Proficient Mean is a statewide statistic comprised of the average or mean score attained on each cluster by all students (GE, SE, and LEP) with a scale score of 200, i.e., students who are "just proficient." Braille students, students taking a breach form and students whose NJ ASK test booklets were coded as "void" were excluded from these means.

The ISR for NJ ASK4 is shown in sample format as Figure 10.1.1 (front page) and Figure 10.1.2 (back page).

Figure 10.1.1

2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Individual Student Report (ISR) – Front

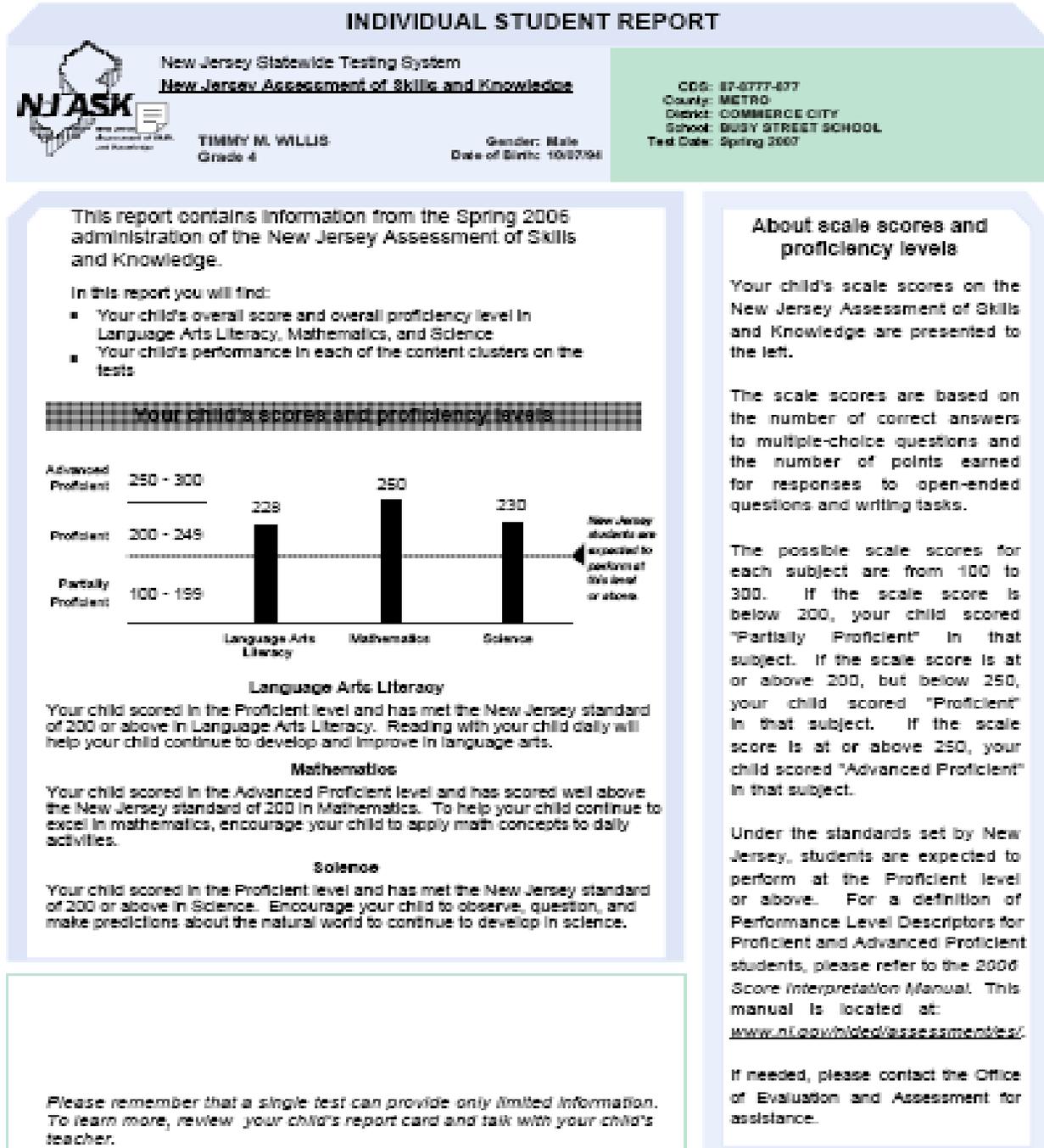


Figure 10.1.2

2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
Individual Student Report (ISR) – Back

More about your child's performance

Language Arts Literacy: Cluster Scores

	Your Child's Points	Total Points Possible	Just Proficient Mean (JPM)
Total	25.0	43.0	19.0
Writing	9.0	20.0	9.5
Writing about Pictures	4.0	10.0	NA
Writing about Poems	5.0	10.0	NA
Reading	16.0	23.0	9.5
Working with Text	7.0	8.0	2.7
Analyzing Text	9.0	15.0	6.8

The Language Arts Literacy test has two clusters: Writing and Reading. The Writing cluster consists of one writing task about a picture and one writing task about a poem. The Reading subcluster, Working with Text, refers to questions that measure students' understanding of ideas and information in the text. The Reading subcluster, Analyzing Text, focuses on the student's analysis of what he/she has read. If your child did not receive a score for a writing task, a code may appear. Possible codes are WF - Wrong Format, NE - Not English, NR - No Response, and OT - Off Topic. Just Proficient Means are not provided for writing tasks.

Mathematics: Cluster Scores

	Your Child's Points	Total Points Possible	Just Proficient Mean (JPM)
Total	33.0	43.0	17.5
Number and Numerical Operations	7.5	13.0	6.0
Geometry and Measurement	8.0	10.0	4.8
Patterns and Algebra	7.5	10.0	3.3
Data Analysis, Probability, and Discrete Mathematics	10.0	10.0	3.5
Problem Solving	24.5	29.0	9.4

The Mathematics test has five clusters. The first four are content clusters and the last is a process cluster. The content clusters refer to four different mathematical topics. The process cluster refers to test questions that are designed to measure mathematical problem-solving ability. Each test question measures one content cluster and may contribute to the process cluster.

Science: Cluster Scores

	Your Child's Points	Total Points Possible	Just Proficient Mean (JPM)
Total	25.5	39.0	19.0
Life Science	10.5	15.0	6.5
Physical Science	8.0	12.0	6.6
Earth Science	7.0	12.0	5.9
Application	20.5	31.0	16.4

The Science test consists of four reporting clusters. The first three - Life, Physical, and Earth - are content clusters. The fourth reporting cluster, Application, refers to test questions that are designed to measure a student's ability to apply skills and knowledge. Each question addresses one content cluster and may contribute to the Application cluster.

About Cluster Scores

A 'cluster' is a group of related test questions on a single topic. The tables on this page show how your child performed on different clusters in Language Arts Literacy, Mathematics, and Science. For each cluster, the tables show the number of points earned by your child, the total number of points possible, and the 'Just Proficient Means' for that cluster.

About Just Proficient Means

The Just Proficient Mean (JPM) is the average points earned in that cluster by all New Jersey students who received a scale score of 200 on the test as a whole. The tables on this page show how your child performed in each cluster relative to the total points possible and Just Proficient students.

A Note on Using this Information

Cluster scores must be interpreted with caution because they are based on a small number of test questions. Students who score about the same as or below the JPM may benefit from additional instruction. When you decide whether your child needs special help in a cluster, you should also take into account your child's performance on classroom work and other activities.

All Sections Roster

The All Sections Roster provides a convenient method for reviewing students' complete test results. The report displays student names in alphabetical order (last name first). Users of this report can quickly determine how a particular student performed in all content areas: Language Arts Literacy, Mathematics and Science.

Following a student's identification information, the student's Scale Score and Proficiency Level (Partially Proficient, Proficient, or Advanced Proficient) are printed for each test section. If the student's test booklet was coded void, the reason code will appear in this space.

Student Roster – Language Arts Literacy

The Student Roster – Language Arts Literacy lists the names of the students (last name first) in groups by proficiency level. Thus, the first students listed on the Language Arts Literacy roster are the students with the highest Language Arts Literacy scale scores. Students are listed alphabetically when more than one student has earned the same score. Students whose test booklets were voided and students coded APA, who are exempt from taking the test, are listed alphabetically at the end of the roster.

Following a student's identification information, the student's Language Arts Literacy scale score is given. This score is based on a combination of the number of correct answers to multiple-choice items and the number of points earned for open-ended items and writing tasks. Points earned are then reported for each cluster. Each item contributes only once to the NJ ASK total score.

Student Roster – Mathematics

The Student Roster – Mathematics lists the names of the students (last name first) in groups by proficiency level. Thus, the first students listed on the Mathematics roster are the students with the highest Mathematics scale scores. Students are listed alphabetically when more than one student has achieved the same score. Students whose test booklets were voided and students coded APA, who are exempt from taking the test, are listed alphabetically at the end of the roster.

Following a student's identification information, the student's total Mathematics score is given. This score is based on a combination of the number of correct answers to multiple-choice items and the number of points earned for open-ended items. Points earned are then reported for each cluster. Each item contributes only once to the NJ ASK total score.

Student Roster – Science

The Student Roster – Science lists the names of the students (last name first) in groups by proficiency level. Thus, the first students listed on the Science roster are the students with the highest Science scale scores. Students are listed alphabetically when more than one student has

achieved the same score. Students whose test booklets were voided and students coded APA, who are exempt from taking the test, are listed alphabetically at the end of the roster.

Following a student's identification information, the student's total Science score is given. This score is based on a combination of the number of correct answers to multiple-choice items and the number of points earned for open-ended items. Points earned are then reported for each cluster. Each item contributes only once to the NJ ASK total score.

Summary of School Performance

There are three Summary of School Performance reports, one for each content area: Language Arts Literacy, Mathematics and Science. The reports are produced at the school level and provide preliminary aggregated data for a test section. Final aggregated data is sent in Cycle II. Data are provided for total students, general education students, special education students, and limited English proficient students. Data are also presented in the report by gender, ethnicity, economic status, and migrant status.

The report provides the percent of students in each proficiency level as well as the number of total students, general education students, special education students, and limited English proficient students tested for each content area.

Summary of District Performance

There are three Summary of District Performance reports, one for each content area: Language Arts Literacy, Mathematics and Science. This report provides aggregated data for the district. In addition, this report includes data for total students, general education students, special education students, and limited English proficient students. The report format is the same as the summary of school performance. Any district that chooses to test a student classified Alternate Proficiency Assessment (APA), who is exempt from taking the NJ ASK, will receive score reports for that student, and the scores will be aggregated into the school and district reports.

Summary of School Cluster Performance

There are three Summary of School Cluster Performance reports, one for each content area: Language Arts Literacy, Mathematics and Science. The reports are produced at the school level and provide aggregated data for each test section. Data are provided for general education students, special education students, limited English proficient students, and Title I students. Cluster level means for each of these populations are also presented on this report.

Summary of District Cluster Performance

There are three Summary of District Cluster Performance reports, one for each content area: Language Arts Literacy, Mathematics and Science. These reports provide aggregated data for the district. In addition, these reports include data for total students, general education students, special education students, limited English proficient students, and Title I students. The report format is the same as the summary of school cluster performance. Any district that chooses to

test a student classified Alternate Proficiency Assessment (APA), who is exempt from taking the NJ ASK, will receive score reports for that student, and the scores will be aggregated into the school and district reports.

10.2 Cycle II Reports

The Cycle II reports include the following: School and District Reports, Special Reports, Statewide Report, DFG Reports, Statewide Charter School Report, Special Needs Report, Non-Special Needs Report, and Title I Report. Each Cycle II report is briefly described below.

School and District Reports

The school and district reports provide a complete analysis of student performance. Separate reports are produced for each subject tested. Each report covers two pages. The first page of each report provides information pertaining to total students, general education students, special education students, and limited English proficient students, as well as to groups classified by gender, ethnicity, economic status, and migrant status. The second page provides cluster raw score information for total students, general education students, special education students, limited English proficient students, and Title I students. This page also contains mean scores for the school or district, for the DFG in which the school or district is classified, and for the State.

For districts and schools identified as “Special Needs”, the Special Needs District Mean is also included. This is the mean as calculated for total students, statewide, in districts identified as “Special Needs.”

The School Report for NJ ASK3 is shown in sample format as Figure 10.2.1 (front page – Performance by Demographic Groups) and Figure 10.2.2 (back page – Summary of Cluster Performance).

Figure 10.2.1

2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
School Report - Performance by Demographic Groups



New Jersey Statewide Testing System
New Jersey Assessment of Skills and Knowledge
Summary of School Performance: Language Arts Literacy
Cycle II Report
Grade 3
Special Needs School

CDS: 99-9999-999
County: MARBLE
District: GRANITE TOWNSHIP
School: ROCKY VALLEY SCHOOL
Test Date: Spring 2007
Page: 1 of 2

	Students Enrolled ¹	APA Students ²	Not Present	Voids	Valid Scale Scores	Partially Proficient		Proficient		Advanced Proficient		Scale Score Mean
						Number	Percent	Number	Percent	Number	Percent	
Total Students³	72	2	1	0	69	18	26.1	50	72.5	1	1.4	209.6
General Education⁴	59	0	0	0	59	15	25.4	43	72.9	1	1.7	211.5
Special Education	7	2	0	0	5	2	40.0	3	60.0	0	0.0	195.0
Limited English Proficient⁵	11	1	1	0	9	2	22.2	7	77.8	0	0.0	204.1
Current LEP	7	1	1	0	5	1	20.0	4	80.0	0	0.0	202.4
Former LEP	4	0	0	0	4	1	25.0	3	75.0	0	0.0	206.3
Gender⁶												
Female	38	2	1	0	35	8	22.9	26	74.3	1	2.9	213.2
Male	34	0	0	0	34	10	29.4	24	70.6	0	0.0	205.9
Ethnicity												
American Indian	0	0	0	0	0	-	-	-	-	-	-	-
Asian	1	0	1	0	0	-	-	-	-	-	-	-
Black	47	1	0	0	46	12	26.1	33	71.7	1	2.2	210.5
Hispanic	23	1	0	0	22	5	22.7	17	77.3	0	0.0	209.0
Pacific Islander	0	0	0	0	0	-	-	-	-	-	-	-
White	1	0	0	0	1	1	100.0	0	0.0	0	0.0	182.0
Other ⁷	0	0	0	0	0	-	-	-	-	-	-	-
Economic Status												
Economically Disadvantaged	41	1	0	0	40	10	25.0	30	75.0	0	0.0	208.6
Non-Economically Disadvantaged	31	1	1	0	29	8	27.6	20	69.0	1	3.4	211.0
Migrant Status												
Migrant	0	0	0	0	0	-	-	-	-	-	-	-
Non-Migrant	72	2	1	0	69	18	26.1	50	72.5	1	1.4	209.6

NOTE: Percentages may not total 100 due to rounding.
- No students in this category.

- ¹ Enrollment is based on the number of scannable test booklets.
- ² Includes Special Education students who did not take the NJ ASK and are required to take the Alternate Proficiency Assessment (APA).
- ³ Students appear in each applicable category; they are included in Total Students only once.

- ⁴ Includes students who were coded Former Limited English Proficient.
- ⁵ Includes Current and Former Limited English Proficient Students.
- ⁶ Excludes students who did not have gender coded.
- ⁷ Includes students who did not have ethnically coded and students who had more than one ethnicity coded.

See page 2
for Cluster Information



Figure 10.2.2

2007 New Jersey Assessment of Skills and Knowledge (NJ ASK)
 School Report – Summary of School Cluster Performance

 New Jersey Statewide Testing System New Jersey Assessment of Skills and Knowledge Summary of School Cluster Performance: Science Cycle I Report (Preliminary Data) Grade 4		CDS: 87-8777-877 County: METRO District: COMMERCE CITY School: BUSY STREET SCHOOL Test Date: Spring 2007 Page: 2 of 2			
CLUSTER MEANS¹					
	Life Science	Physical Science	Earth Science	Application	Total
Total Points Possible	15.0	12.0	12.0	31.0	39.0
Just Proficient Means²	6.5	6.6	5.9	16.4	19.0
Total Students³	8.7	7.5	6.8	19.8	23.0
General Education⁴	9.0	7.9	7.2	20.7	24.1
Special Education	7.3	5.7	4.9	15.8	17.9
Limited English Proficient⁵	7.0	7.0	5.0	16.0	19.0
Current LEP	7.0	7.0	5.0	16.0	19.0
Former LEP	0.0	0.0	0.0	0.0	0.0
Title I	0.0	0.0	0.0	0.0	0.0
School Mean	8.7	7.5	6.8	19.8	23.0
District Mean	8.8	7.2	7.3	19.7	23.3
Students coded both SE and Current LEP 0					
<small> - No Students in this category. 1 Cluster means exclude students who took alternate test forms, as well as students with voided test booklets. 2 Just Proficient Means are the statewide raw score means for students whose scale scores are 200. 3 Students appear in each applicable category; they are included in Total Students only once. 4 Includes students who were coded Former Limited English Proficient. 5 Includes Current and Former Limited English Proficient Students. </small>					

Special Reports

Special reports are produced when a district requests information about the performance of special groups, as identified by the district at the time of testing. By using the “special” code category at the time of the test administration, districts have the opportunity to create such reports for specific student groups containing six or more students. Student test booklets may be coded in any of the four two-column “Special Codes” grids labeled A, B, C, and D. The special code, as coded on the students’ test booklet, is printed in the report title. These reports are produced at the school level. One report for each content area per code is produced.

Statewide Report

The Statewide Report provides state-level data pertaining to the performance of the total student population, as well as the general education, special education and limited English proficient students. Performance is also reported by gender, ethnicity, economic status and migrant status.

District Factor Group (DFG) Report

The DFG Report summarizes the performance data for each DFG by total students, general education students, special education students and limited English proficient students as well as for groups classified by gender, ethnicity, economic status and migrant status. There is one DFG report for each District Factor Group.

Charter School Report

The Charter School Report summarize the performance data by total students, general education students, special education students and limited English proficient students as well as for groups classified by gender, ethnicity, economic status and migrant status, for all students in charter schools within the State.

Special Needs Report

The Special Needs Report summarize the statewide performance of students in special needs districts. Results are reported by total students, general education students, special education students and limited English proficient students as well as for student groups classified by gender, ethnicity, economic status and migrant status.

Non-Special Needs Report

The Non-Special Needs Reports summarize the statewide performance of student in districts not designated as special needs districts. Results are reported by total students, general education students, special education students and limited English proficient students as well as for student groups classified by gender, ethnicity, economic status and migrant status.

Title I Report

The Title I Reports summarize the performance of Title I students statewide. Results are reported by total students, general education students, special education students and limited English proficient students as well as for student groups classified by gender, ethnicity, economic status and migrant status.

10.3 State Summary Reporting

The State Summary consists of a group of files presented to the State on a CD. These files include an executive summary, report PDFs, and test result tables and graphs.

The executive summary contains a brief history of each test and the highlights of 2007 results based on the state Cycle II demographic report. The executive summaries for Grades 3 and 4 can be found in Appendix A. Additional statewide Cycle II results can be found in Appendix B. Two files of test results are provided based on the Cycle II demographic reports; one file includes all data with no suppression rules applied, and the other file applies the suppression rules for small cell numbers. The suppression rules are included in the executive summaries in Appendix A.

The report PDFs included on the State Summary CD are the DFG Reports, the Charter School Reports, the Special Needs Reports, the Non-Special Needs Reports and the Statewide Reports.

Longitudinal data graphs of percentages proficient and above for demographic groups from the first time each test was administered to 2007 are also provided in the State Summary.

10.4 Interpreting Reports

The 2007 NJ ASK score report information is used for the purpose of district monitoring. The data are also provided to assist districts in the review of current curricular programs. With the adoption of the Core Curriculum Content Standards in May 1996, all districts were required to implement standards based instruction. NJ ASK results displayed in school-level and district-level reports can provide meaningful information for educational program reviews.

All other factors being equal, the reliability (stability) of scores decreases as the number of items used decreases. Generally speaking, reliability is lower in clusters that have smaller numbers of items. All else being equal, differences in mean cluster scores for clusters with smaller numbers of items must be greater than differences for clusters with large numbers of items before they can be considered meaningful. Decreases in reliability also increase the need for multiple measures, particularly where the number of students in the assessed group is small.

All clusters cannot be assumed to be of equal difficulty level. Cluster scores should, therefore, be compared to their respective Just Proficient Means to facilitate effective interpretation. Insofar as tests are not equated at the cluster level, cluster scores cannot be compared from year to year. Year-to-year comparisons should be limited to total test scores in the subjects tested. For each subject, it is the whole test level (only) for which scores are equated.

The NJ ASK reports provide information on clusters in content areas that need further attention. However, since some clusters were assessed with a relatively small number of items, evaluation of a student's performance should never be based solely on the results of the NJ ASK or any other single form of formal or informal assessment. Insofar as the NJ ASK is equated at the test level only, cluster performance should not be directly compared across multiple test administrations.

10.5 Quality Control in Reporting

Prior to reports being distributed, both the reports themselves and the steps leading up to the production of the reports are subjected to extensive quality control procedures. These procedures include tasks to ensure the raw scores are accurately recorded in the database, and to ensure the scale scores and proficiency levels have been converted accurately. The aggregated data file is extensively reviewed to ensure the data are aggregated according to the aggregation rules defined by the State. The paper reports are then reviewed to verify all of the data is accurately represented on each report.

APPENDIX A: Statewide Cycle II Executive Summary Results

Grade 3 New Jersey Assessment of Knowledge and Skills Spring 2007

The spring 2007 grade 3 New Jersey Assessment of Knowledge and Skills (NJ ASK) consisted of two content areas: Language Arts Literacy and Mathematics. The NJ ASK is designed to give an early indication of the progress students are making in mastering the knowledge and skills described in the Core Curriculum Content Standards. The results are to be used by schools and districts to identify strengths and weaknesses in their educational programs. It is anticipated that this process will lead to improved instruction and better alignment with the Core Curriculum Content Standards. The results may also be used, along with other indicators of student progress, to identify those students who may need instructional support in any of the content areas. This support, which could be in the form of individual or programmatic intervention, would be a means to address any identified knowledge or skill gaps.

The NJ ASK scores are reported as scale scores in each of the content areas. The scores range from 100-199 (Partially Proficient), 200-249 (Proficient), and 250-300 (Advanced Proficient). The scores of students who are included in the Partially Proficient level are considered to be below the state minimum of proficiency and those students may be most in need of instructional support.

The NJ ASK was administered in March 2007. From a total third grade student population of 102,812, valid scores were obtained in Language Arts Literacy from 100,877 students, with 289 students not present and 997 voids (unscorable due to illness, other difficulties during testing, or an insufficient number of items answered in a given content area). Valid scores were obtained in Mathematics from 101,800 students, with 238 not present and 163 voids.

This executive summary includes two tables summarizing statewide test results for the 2007 grade 3 administration of the NJ ASK. Table A.3.1 presents results for Language Arts Literacy and Table A.3.2 presents results for Mathematics. Results are presented for the following student groups: all, special education, and limited English proficient (LEP) students. LEP is further broken out by the following groups: LEP current and former, LEP current, and LEP former. Data are also summarized for several demographic variables including: gender, ethnicity, and economic status. The tables include the number of students enrolled, not present, voided, and with valid scale scores. In addition, the tables present mean scale score and the percent of students in each performance category (i.e., Partially Proficient, Proficient, and Advanced Proficient).

The tables that follow are derived from the statewide performance data of the Cycle II report. Note that the enrollment is based on the number of students with scannable test booklets. Also, students coded as multiple ethnicity and those whose ethnicity was unspecified are counted as Other. The percentage of students in the combined category, Proficient or Advanced Proficient, is calculated by subtracting the percentage of students in Partially Proficient from one hundred. The percentages may not total to one hundred due to rounding.

Following are highlights of the 2007 third grade assessment results.

Grade 3 Language Arts Literacy Results:

- Of the 100,877 grade 3 students with valid scale scores in Language Arts Literacy in spring 2007, 16.6% scored in Partially Proficient, 75.2% scored in Proficient and 8.2% scored in Advanced Proficient (Table A.3.1).
- **Special Education** 56.8% of special education grade 3 students in 2007 scored in Proficient or Advanced Proficient in Language Arts Literacy (Table A.3.1).
- **Limited English Proficient, Current and Former** 61.4% of total limited English proficient grade 3 students in 2007 scored in Proficient or Advanced Proficient in Language Arts Literacy (Table A.3.1).
- **Current Limited English Proficient** 51.3% of current limited English proficient grade 3 students in 2007 scored in Proficient or Advanced Proficient in Language Arts Literacy (Table A.3.1).
- **Former Limited English Proficient** 77.5% of former limited English proficient grade 3 students in 2007 scored in Proficient or Advanced Proficient in Language Arts Literacy (Table A.3.1).
- **Gender** 86.6% of female compared to 80.4% of male grade 3 students in 2007 scored in Proficient or Advanced Proficient in Language Arts Literacy (Table A.3.1).
- **Ethnicity** For performance by grade 3 ethnic groups in 2007, students scoring in Proficient or Advanced Proficient in Language Arts Literacy ranged from 92.7% of Asian students to 68.7% of African American students. The percentage of Proficient and Advanced Proficient for all other race/ethnic groups fell between Asians and African Americans (Table A.3.1).
- **Economic Status** 69.0% of economically disadvantaged grade 3 students in 2007 scored in Proficient or Advanced Proficient in Language Arts Literacy (Table A.3.1).
- The mean scale score for all grade 3 students on the Language Arts Literacy test in spring 2007 was 218.0 (Table A.3.1).

Grade 3 Mathematics Results:

- Of the 102,812 grade 3 students with valid scale scores in Mathematics in spring 2007, 12.7% scored in Partially Proficient, 55.0% scored in Proficient and 32.3% scored in Advanced Proficient (Table A.3.2).
- **Special Education** 71.9% of special education grade 3 students in 2007 scored in Proficient or Advanced Proficient in Mathematics (Table A.3.2).

- **Limited English Proficient, Current and Former** 71.6% of total limited English proficient grade 3 students in 2007 scored in Proficient or Advanced Proficient in Mathematics (Table A.3.2).
- **Current Limited English Proficient** 65.1% of current limited English proficient grade 3 students in 2007 scored in Proficient or Advanced Proficient in Mathematics (Table A.3.2).
- **Former Limited English Proficient** 84.2% of former limited English proficient grade 3 students in 2007 scored in Proficient or Advanced Proficient in Mathematics (Table A.3.2).
- **Gender** 87.7% of female compared to 87.0% of male grade 3 students in 2007 scored in Proficient or Advanced Proficient in Mathematics (Table A.3.2).
- **Ethnicity** For performance by grade 3 ethnic groups in 2007, students scoring in Proficient or Advanced Proficient in Mathematics ranged from 96.1% of Asian students to 72.6% of African American students. The percentage of Proficient and Advanced Proficient for all other race/ethnic groups fell between Asians and African Americans (Table A.3.2).
- **Economic Status** 75.6% of economically disadvantaged grade 3 students in 2007 scored in Proficient or Advanced Proficient in Mathematics (Table A.3.2).
- The mean scale score for all grade 3 students on the Mathematics test in spring 2007 was 232.3 (Table A.3.2).

Reporting Rules for Data File:

The accompanying state summary data file contains the same type of information shown in the statewide summary tables included with this executive summary. Please note that there may be small differences between the state summary data file and the Cycle II reports issued to districts. In order to safeguard student confidentiality, certain information is suppressed in the state summary file according to the following reporting rules:

- Data are not reported where the number of students with valid scale scores for a particular group is less than 11.
- Data are not reported where demographic groups are mutually exclusive (e.g., gender) and there are one or two students with a valid scale score in one of the groups (e.g., male).
- Data are not reported when it is otherwise possible to identify individual student performance.

TABLE A.3.1

STATEWIDE PERFORMANCE BY DEMOGRAPHIC GROUPS SPRING 2007 -- GRADE 3

Language Arts Literacy

		Number of Students Enrolled	Number Not Present	Number of Voids	Number of Valid Scale Scores	Scale Score Mean	Percent Partially Proficient	Percent Proficient	Percent Advanced Proficient
Total	All Students	102812	289	997	100877	218.0	16.6%	75.2%	8.2%
Education Status	Special Education Students	15275	75	184	14367	199.1	43.2%	55.0%	1.9%
Limited English Proficient	Limited English Proficient Students (Current and Former)	6626	85	654	5874	202.0	38.6%	59.2%	2.2%
	Current Limited English Proficient Students	4363	80	651	3625	196.0	48.7%	50.0%	1.4%
	Former Limited English Proficient Students	2263	5	3	2249	211.7	22.5%	74.0%	3.5%
Gender	Female	49756	112	391	49044	221.5	13.4%	75.2%	11.4%
	Male	53007	166	604	51808	214.6	19.6%	75.3%	5.0%
Ethnicity	American Indian	76	0	1	75	212.7	25.3%	66.7%	8.0%
	Asian	8514	27	148	8287	227.6	7.3%	75.7%	17.0%
	Black	17722	66	175	17354	206.7	31.3%	65.9%	2.7%
	Hispanic	19590	102	424	18926	208.5	27.8%	69.2%	2.9%
	Pacific Islander	211	1	3	206	224.4	7.8%	78.2%	14.1%
	White	56143	63	236	55534	223.3	9.6%	80.1%	10.3%
	Other	556	30	10	495	214.8	21.2%	71.7%	7.1%
Economic Status	Economically Disadvantaged	31727	127	547	30829	206.5	31.0%	66.6%	2.4%
	Non-Economically Disadvantaged	71085	162	450	70048	223.0	10.3%	79.0%	10.7%

TABLE A.3.2

STATEWIDE PERFORMANCE BY DEMOGRAPHIC GROUPS SPRING 2007 -- GRADE 3

Mathematics

		Number of Students Enrolled	Number Not Present	Number of Voids	Number of Valid Scale Scores	Scale Score Mean	Percent Partially Proficient	Percent Proficient	Percent Advanced Proficient
Total	All Students	102812	238	163	101800	232.3	12.7%	55.0%	32.3%
Education Status	Special Education Students	15275	67	89	14508	217.2	28.1%	53.1%	18.8%
Limited English Proficient	Limited English Proficient Students (Current and Former)	6626	26	8	6582	216.1	28.4%	55.8%	15.8%
	Current Limited English Proficient Students	4363	22	7	4330	210.5	34.9%	53.1%	12.0%
	Former Limited English Proficient Students	2263	4	1	2252	226.9	15.8%	61.0%	23.2%
Gender	Female	49756	91	48	49414	232.2	12.3%	56.3%	31.4%
	Male	53007	143	114	52353	232.5	13.0%	53.9%	33.2%
Ethnicity	American Indian	76	0	0	76	227.1	15.8%	61.8%	22.4%
	Asian	8514	16	8	8439	246.3	3.9%	42.3%	53.8%
	Black	17722	82	62	17456	215.6	27.4%	58.6%	14.0%
	Hispanic	19590	61	28	19374	221.5	21.1%	60.7%	18.1%
	Pacific Islander	211	1	0	209	240.2	7.2%	50.2%	42.6%
	White	56143	60	61	55733	239.2	6.4%	53.9%	39.7%
	Other	556	18	4	513	228.3	17.9%	52.8%	29.2%
Economic Status	Economically Disadvantaged	31727	118	76	31324	218.5	24.4%	59.5%	16.2%
	Non-Economically Disadvantaged	71085	120	87	70476	238.5	7.5%	53.1%	39.5%

Statewide Cycle II Executive Summary Results

Grade 4 New Jersey Assessment of Knowledge and Skills Spring 2007

The spring 2007 grade 4 New Jersey Assessment of Knowledge and Skills (NJ ASK) consisted of three content areas: Language Arts Literacy, Mathematics, and Science. The NJ ASK is designed to give an early indication of the progress students are making in mastering the knowledge and skills described in the Core Curriculum Content Standards. The results are to be used by schools and districts to identify strengths and weaknesses in their educational programs. It is anticipated that this process will lead to improved instruction and better alignment with the Core Curriculum Content Standards. The results may also be used, along with other indicators of student progress, to identify those students who may need instructional support in any of the content areas. This support, which could be in the form of individual or programmatic intervention, would be a means to address any identified knowledge or skill gaps.

The NJ ASK scores are reported as scale scores in each of the content areas. The scores range from 100-199 (Partially Proficient), 200-249 (Proficient), and 250-300 (Advanced Proficient). The scores of students who are included in the Partially Proficient level are considered to be below the state minimum of proficiency, and those students may be most in need of instructional support.

The NJ ASK was administered in March 2007. From a total fourth grade student population of 102,490, valid scores were obtained in Language Arts Literacy from 100,617 students, with 339 students not present and 879 voids (unscorable due to illness, other difficulties during testing, or an insufficient number of items answered in a given content area). Valid scores were obtained in Mathematics from 101,310 students, with 278 not present and 259 voids. Valid scores were obtained in Science from 101,266 students, with 497 not present and 130 voids.

This executive summary includes three tables summarizing statewide test results for the 2007 administration of the grade 4 NJ ASK. Table A.4.1 presents results for Language Arts Literacy, Table A.4.2 presents results for Mathematics, and Table A.4.3 presents results for Science. Results are presented for the following student groups: all, special education, and limited English proficient (LEP) students. LEP is further broken out by the following groups: LEP current and former, LEP current, and LEP former. Data are also summarized for several demographic variables including: gender, ethnicity, and economic status. The tables include the number of students enrolled, not present, voided, and with valid scale scores. In addition, the tables present mean scale score and the percent of students in each performance category (i.e., Partially Proficient, Proficient, and Advanced Proficient).

The tables that follow are derived from the statewide performance data of the Cycle II report. Note that the enrollment is based on the number of students with scannable test booklets. Also, students coded as multiple ethnicity and those whose ethnicity was unspecified are counted as Other. The percentage of students in the combined category, Proficient or Advanced Proficient, is calculated by subtracting the percentage of students in Partially Proficient from one hundred. The percentages may not total to one hundred due to rounding.

Following are highlights of the 2007 fourth grade assessment results.

Grade 4 Language Arts Literacy Results:

- Of the 100,617 grade 4 students with valid scale scores in Language Arts Literacy in spring 2007, 19.4% scored in Partially Proficient, 73.9% scored in Proficient and 6.7% scored in Advanced Proficient (Table A.4.1).
- **Special Education** 49.5% of special education grade 4 students in 2007 scored in Proficient or Advanced Proficient in Language Arts Literacy (Table A.4.1).
- **Limited English Proficient, Current and Former** 53.3% of total limited English proficient grade 4 students in 2007 scored in Proficient or Advanced Proficient in Language Arts Literacy (Table A.4.1).
- **Current Limited English Proficient** 44.6% of current limited English proficient grade 4 students in 2007 scored in Proficient or Advanced Proficient in Language Arts Literacy (Table A.4.1).
- **Former Limited English Proficient** 68.4% of former limited English proficient grade 4 students in 2007 scored in Proficient or Advanced Proficient in Language Arts Literacy (Table A.4.1).
- **Gender** 84.3% of female compared to 77.1% of male grade 4 students in 2007 scored in Proficient or Advanced Proficient in Language Arts Literacy (Table A.4.1).
- **Ethnicity** For performance by grade 4 ethnic groups in 2007, students scoring in Proficient or Advanced Proficient in Language Arts Literacy ranged from 91.8% of Asian students to 62.9% of African American students. The percentage of Advanced Proficient and Proficient for all other race/ethnic groups fell between Asians and African Americans (Table A.4.1).
- **Economic Status** 64.5% of economically disadvantaged grade 4 students in 2007 scored in Proficient or Advanced Proficient in Language Arts Literacy (Table A.4.1).
- The mean scale score for all grade 4 students on the Language Arts Literacy test in spring 2007, was 215.7 (Table A.4.1).

Grade 4 Mathematics Results:

- Of the 101,310 grade 4 students with valid scale scores in Mathematics in spring 2007, 15.3% scored in Partially Proficient, 43.7% scored in Proficient and 41.0% scored in Advanced Proficient (Table A.4.2).

- **Special Education** 63.7% of special education grade 4 students in 2007 scored in Proficient or Advanced Proficient in Mathematics (Table A.4.2).
- **Limited English Proficient, Current and Former** 62.2% of total limited English proficient grade 4 students in 2007 scored in Proficient or Advanced Proficient in Mathematics (Table A.4.2).
- **Current Limited English Proficient** 54.8% of current limited English proficient grade 4 students in 2007 scored in Proficient or Advanced Proficient in Mathematics (Table A.4.2).
- **Former Limited English Proficient** 77.7% of former limited English proficient grade 4 students in 2007 scored in Proficient or Advanced Proficient in Mathematics (Table A.4.2).
- **Gender** 84.8% of female compared to 84.6% of male grade 4 students in 2007 scored in Proficient or Advanced Proficient in Mathematics (Table A.4.2).
- **Ethnicity** For performance by grade 4 ethnic groups in 2007, students scoring in Proficient or Advanced Proficient in Mathematics ranged from 94.9% of Asian students to 68.3% of African American students. The percentage of Advanced Proficient and Proficient for all other race/ethnic groups fell between Asians and African Americans (Table A.4.2).
- **Economic Status** 71.3% of economically disadvantaged grade 4 students in 2007 scored in Proficient or Advanced Proficient in Mathematics (Table A.4.2).
- The mean scale score for all grade 4 students on the Mathematics test in spring 2007 was 234.1 (Table A.4.2).

Grade 4 Science Results:

- Of the 101,266 grade 4 students with valid scale scores in Science in spring 2007, 16.9% scored in Partially Proficient, 42.3% scored in Proficient and 40.8% scored in Advanced Proficient (Table A.4.3).
- **Special Education** 66.4% of special education grade 4 students in 2007 scored in Proficient or Advanced Proficient in Science (Table A.4.3).
- **Limited English Proficient, Current and Former** 50.4% of total limited English proficient grade 4 students in 2007 scored in Proficient or Advanced Proficient in Science (Table A.4.3).
- **Current Limited English Proficient** 41.3% of current limited English proficient grade 4 students in 2007 scored in Proficient or Advanced Proficient in Science (Table A.4.3).

- **Former Limited English Proficient** 69.4% of former limited English proficient grade 4 students in 2007 scored in Proficient or Advanced Proficient in Science (Table A.4.3).
- **Gender** 82.5% of female compared to 83.8% of male grade 4 students in 2007 scored in Proficient or Advanced Proficient in Science (Table A.4.3).
- **Ethnicity** For performance by grade 4 ethnic groups in 2007, students scoring in Proficient or Advanced Proficient in Science ranged from 92.4% of White students to 64.5% of African American students. The percentage of Advanced Proficient and Proficient for all other race/ethnic groups fell between Asians and African Americans (Table A.4.3).
- **Economic Status** 66.2% of economically disadvantaged grade 4 students in 2007 scored in Proficient or Advanced Proficient in Science (Table A.4.3).
- The mean scale score for all grade 4 students on the Science test in spring 2007 was 231.3 (Table A.4.3).

Reporting Rules for Data File

The accompanying state summary data file contains the same type of information shown in the statewide summary tables included with this executive summary. Please note that there may be small differences between the data file and reports issued to districts due to adjustments made to the data by districts after the reporting deadline. In order to safeguard student confidentiality, certain information is suppressed in the state summary file according to the following reporting rules:

- Data are not reported where the number of students with valid scale scores for a particular group is less than 11.
- Data are not reported where demographic groups are mutually exclusive (e.g., gender) and there are one or two students with a valid scale score in one of the groups (e.g., male).
- Data are not reported when it is otherwise possible to identify individual student performance.

TABLE A.4.1

STATEWIDE PERFORMANCE BY DEMOGRAPHIC GROUPS SPRING 2007 -- GRADE 4

Language Arts Literacy

		Number of Students Enrolled	Number Not Present	Number of Voids	Number of Valid Scale Scores	Scale Score Mean	Percent Partially Proficient	Percent Proficient	Percent Advanced Proficient
Total	All Students	102490	339	879	100617	215.7	19.4%	73.9%	6.7%
Education Status	Special Education Students	16130	83	163	15229	194.1	50.5%	48.3%	1.2%
Limited English Proficient	Limited English Proficient Students (Current and Former)	5521	89	589	4830	196.4	46.7%	52.0%	1.3%
	Current Limited English Proficient Students	3747	89	586	3063	190.5	55.4%	43.9%	0.7%
	Former Limited English Proficient Students	1774	0	3	1767	206.5	31.6%	65.9%	2.4%
Gender	Female	49817	148	359	49096	219.3	15.7%	74.9%	9.4%
	Male	52633	184	518	51502	212.2	22.9%	73.0%	4.1%
Ethnicity	American Indian	111	0	1	108	209.3	25.9%	70.4%	3.7%
	Asian	8304	27	110	8119	226.8	8.2%	76.3%	15.6%
	Black	17495	81	152	17131	202.9	37.1%	60.9%	2.0%
	Hispanic	19277	104	414	18642	206.3	30.9%	66.7%	2.4%
	Pacific Islander	230	1	5	223	224.2	12.6%	72.6%	14.8%
	White	56561	93	186	55965	221.1	11.8%	80.0%	8.2%
	Other	512	33	11	429	209.8	25.9%	68.8%	5.4%
Economic Status	Economically Disadvantaged	30922	160	507	30050	203.4	35.5%	62.8%	1.7%
	Non-Economically Disadvantaged	71568	179	372	70567	220.9	12.6%	78.6%	8.8%

TABLE A.4.2

STATEWIDE PERFORMANCE BY DEMOGRAPHIC GROUPS SPRING 2007 -- GRADE 4

Mathematics

		Number of Students Enrolled	Number Not Present	Number of Voids	Number of Valid Scale Scores	Scale Score Mean	Percent Partially Proficient	Percent Proficient	Percent Advanced Proficient
Total	All Students	102490	278	259	101310	234.1	15.3%	43.7%	41.0%
Education Status	Special Education Students	16130	92	109	15286	212.2	36.3%	43.1%	20.6%
Limited English Proficient	Limited English Proficient Students (Current and Former)	5521	22	22	5465	210.8	37.8%	43.7%	18.5%
	Current Limited English Proficient Students	3747	19	19	3701	204.7	45.2%	40.2%	14.6%
	Former Limited English Proficient Students	1774	3	3	1764	223.7	22.3%	51.1%	26.6%
Gender	Female	49817	108	83	49411	233.4	15.2%	45.5%	39.3%
	Male	52633	162	174	51881	234.7	15.4%	41.9%	42.6%
Ethnicity	American Indian	111	0	0	109	226.3	19.3%	49.5%	31.2%
	Asian	8304	14	6	8239	251.6	5.1%	29.0%	65.8%
	Black	17495	88	103	17172	214.8	31.7%	49.0%	19.2%
	Hispanic	19277	60	53	19051	221.0	25.3%	49.6%	25.0%
	Pacific Islander	230	0	1	228	246.3	7.0%	32.9%	60.1%
	White	56561	93	89	56068	241.9	8.4%	42.2%	49.5%
	Other	512	23	7	443	227.0	21.9%	43.6%	34.5%
Economic Status	Economically Disadvantaged	30922	130	133	30461	217.7	28.7%	49.4%	21.9%
	Non-Economically Disadvantaged	71568	148	126	70849	241.1	9.6%	41.2%	49.2%

TABLE A.4.3

STATEWIDE PERFORMANCE BY DEMOGRAPHIC GROUPS SPRING 2007 -- GRADE 4

Science

		Number of Students Enrolled	Number Not Present	Number of Voids	Number of Valid Scale Scores	Scale Score Mean	Percent Partially Proficient	Percent Proficient	Percent Advanced Proficient
Total	All Students	102490	497	130	101266	231.3	16.9%	42.3%	40.8%
Education Status	Special Education Students	16130	162	60	15311	213.8	33.6%	44.0%	22.4%
Limited English Proficient	Limited English Proficient Students (Current and Former)	5521	47	9	5454	199.6	49.6%	38.5%	11.9%
	Current Limited English Proficient Students	3747	44	7	3689	193.0	58.7%	32.6%	8.7%
	Former Limited English Proficient Students	1774	3	2	1765	213.6	30.6%	50.8%	18.6%
Gender	Female	49817	217	42	49360	229.6	17.5%	44.8%	37.7%
	Male	52633	274	87	51886	233.0	16.2%	40.0%	43.8%
Ethnicity	American Indian	111	1	0	108	227.3	17.6%	49.1%	33.3%
	Asian	8304	27	5	8231	244.5	8.2%	32.9%	58.9%
	Black	17495	137	47	17183	210.9	35.5%	47.7%	16.9%
	Hispanic	19277	110	14	19046	215.0	30.9%	47.8%	21.4%
	Pacific Islander	230	0	2	227	242.2	9.7%	34.8%	55.5%
	White	56561	198	58	56025	241.3	7.6%	40.3%	52.1%
	Other	512	24	4	446	225.1	23.1%	43.9%	33.0%
Economic Status	Economically Disadvantaged	30922	230	55	30453	212.4	33.8%	47.6%	18.6%
	Non-Economically Disadvantaged	71568	267	75	70813	239.5	9.6%	40.1%	50.3%

**APPENDIX B:
Additional Statewide Cycle II Results**

TABLE B.3.1

**NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007
NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE BY DISTRICT
FACTOR GROUP**

LANGUAGE ARTS LITERACY SECTION – Grade 3

GENERAL EDUCATION STUDENTS ^b

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	13,475	74.2	25.8	71.3	2.9	209.9
B	8,192	83.8	16.2	79.6	4.2	215.9
CD	7,730	87.6	12.4	81.4	6.1	219.1
DE	10,611	91.5	8.5	83.6	8.0	222.9
FG	10,223	93.8	6.2	84.3	9.5	224.5
GH	11,709	94.3	5.7	80.7	13.5	227.0
I	16,107	96.2	3.8	80.7	15.5	229.6
J	3,647	97.9	2.1	77.7	20.3	232.7

SPECIAL EDUCATION STUDENTS ^c

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	2,314	31.1	68.9	30.8	0.3	182.8
B	1,514	44.5	55.5	44.1	0.5	192.0
CD	1,320	46.4	53.6	45.5	0.9	193.2
DE	1,867	61.8	38.2	60.3	1.4	201.4
FG	1,850	63.5	36.5	61.1	2.3	203.5
GH	2,054	65.5	34.5	62.8	2.8	204.9
I	2,742	73.3	26.7	70.0	3.3	209.2
J	529	78.8	21.2	75.0	3.8	212.2

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
d. INCLUDES CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.3.1 (continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM
 SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE
 BY DISTRICT FACTOR GROUP**

LANGUAGE ARTS LITERACY SECTION – Grade 3

CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ^d

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	1,985	48.0	52.0	46.9	1.1	193.8
B	514	48.1	51.9	47.1	1.0	193.8
CD	237	48.5	51.5	47.7	0.8	195.3
DE	182	52.2	47.8	50.5	1.6	197.5
FG	205	61.5	38.5	60.5	1.0	201.8
GH	231	61.0	39.0	58.9	2.2	202.2
I	160	72.5	27.5	67.5	5.0	208.6
J	35	74.3	25.7	68.6	5.7	211.5

TOTAL STUDENTS ^e

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	17,590	66.1	33.9	63.7	2.4	204.8
B	10,193	76.3	23.7	72.8	3.5	211.3
CD	9,266	80.8	19.2	75.6	5.2	214.9
DE	12,636	86.7	13.3	79.8	6.9	219.5
FG	12,255	88.8	11.2	80.5	8.3	221.0
GH	13,968	89.6	10.4	77.8	11.8	223.4
I	18,989	92.7	7.3	79.1	13.7	226.5
J	4,207	95.4	4.6	77.3	18.1	230.0

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
- b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
- c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
- d. INCLUDES CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
- e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.3.1 (continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM
SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE**

LANGUAGE ARTS LITERACY SECTION – Grade 3

CHARTER SCHOOLS^f

	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL ^b EDUCATION STUDENTS	1,152	74.7	25.3	69.6	5.0	211.5
SPECIAL ^c EDUCATION STUDENTS	103	40.8	59.2	37.9	2.9	192.1
CURRENT LIMITED ENGLISH ^d PROFICIENT STUDENTS	48	64.6	35.4	64.6	0.0	204.9
FORMER LIMITED ENGLISH ^d PROFICIENT STUDENTS	0	--	--	--	--	--
TOTAL LIMITED ENGLISH ^d PROFICIENT STUDENTS	48	64.6	35.4	64.6	0.0	204.9
TOTAL ^e STUDENTS	1,299	71.7	28.3	67.1	4.7	209.8

STATEWIDE RESULTS

	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL ^b EDUCATION STUDENTS	83,218	89.1	10.9	79.6	9.5	222.0
SPECIAL ^c EDUCATION STUDENTS	14,367	56.8	43.2	55.0	1.9	199.1
CURRENT LIMITED ENGLISH ^d PROFICIENT STUDENTS	3,625	51.3	48.7	50.0	1.4	196.0
FORMER LIMITED ENGLISH ^d PROFICIENT STUDENTS	2,249	77.5	22.5	74.0	3.5	211.7
TOTAL LIMITED ENGLISH ^d PROFICIENT STUDENTS	5,874	61.4	38.6	59.2	2.2	202.0
TOTAL ^e STUDENTS	100,877	83.4	16.6	75.2	8.2	218.0

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
d. INCLUDES CURRENT, FORMER, OR TOTAL (CURRENT+FORMER) LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
e. INCLUDES ALL STUDENTS TESTED.
f. CHARTER SCHOOLS ARE NOT INCLUDED IN A DFG.
NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.3.2

**NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007
NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE BY DISTRICT
FACTOR GROUP**

MATHEMATICS SECTION – Grade 3

GENERAL EDUCATION STUDENTS ^b

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	13,510	77.8	22.2	59.2	18.6	220.9
B	8,212	86.3	13.7	62.7	23.6	228.3
CD	7,743	89.8	10.2	61.9	27.9	232.3
DE	10,616	93.5	6.5	57.7	35.8	237.7
FG	10,226	94.5	5.5	57.0	37.6	238.9
GH	11,713	96.0	4.0	52.5	43.5	242.3
I	16,118	97.2	2.8	47.8	49.4	245.3
J	3,652	98.5	1.5	42.1	56.4	248.9

SPECIAL EDUCATION STUDENTS ^c

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	2,367	48.5	51.5	40.9	7.6	198.0
B	1,535	65.0	35.0	52.1	12.9	210.3
CD	1,332	64.3	35.7	52.7	11.6	209.6
DE	1,878	74.9	25.1	55.1	19.8	219.8
FG	1,865	77.4	22.6	56.2	21.2	221.6
GH	2,069	80.9	19.1	57.8	23.1	224.8
I	2,750	84.3	15.7	56.7	27.5	228.7
J	532	88.2	11.8	56.4	31.8	231.6

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
- b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
- c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
- d. INCLUDES CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
- e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.3.2 (continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM
 SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE
 BY DISTRICT FACTOR GROUP**

MATHEMATICS SECTION – Grade 3

CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ^d

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	2,232	62.4	37.6	51.3	11.1	208.7
B	655	60.3	39.7	53.0	7.3	205.4
CD	288	66.7	33.3	55.2	11.5	210.3
DE	227	68.3	31.7	57.3	11.0	211.4
FG	258	70.9	29.1	55.8	15.1	216.1
GH	298	70.8	29.2	56.0	14.8	215.1
I	230	80.0	20.0	59.1	20.9	223.1
J	58	81.0	19.0	32.8	48.3	235.5

TOTAL STUDENTS ^e

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	17,912	72.3	27.7	56.0	16.4	216.5
B	10,372	81.6	18.4	60.6	21.1	224.3
CD	9,342	85.6	14.4	60.4	25.2	228.5
DE	12,697	90.4	9.6	57.4	33.0	234.6
FG	12,321	91.5	8.5	56.8	34.7	235.9
GH	14,053	93.3	6.7	53.4	39.9	239.2
I	19,077	95.2	4.8	49.2	46.0	242.6
J	4,238	97.0	3.0	43.8	53.2	246.5

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
- b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
- c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
- d. INCLUDES CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
- e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.3.2 (continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM
SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE**

MATHEMATICS SECTION – Grade 3

CHARTER SCHOOLS^f

	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL ^b EDUCATION STUDENTS	1,157	74.8	25.2	57.7	17.1	218.1
SPECIAL ^c EDUCATION STUDENTS	104	70.2	29.8	53.8	16.3	211.4
CURRENT LIMITED ENGLISH ^d PROFICIENT STUDENTS	47	74.5	25.5	63.8	10.6	216.6
FORMER LIMITED ENGLISH ^d PROFICIENT STUDENTS	0	--	--	--	--	--
TOTAL LIMITED ENGLISH ^d PROFICIENT STUDENTS	47	74.5	25.5	63.8	10.6	216.6
TOTAL ^e STUDENTS	1,304	74.5	25.5	57.7	16.9	217.6

STATEWIDE RESULTS

	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL ^b EDUCATION STUDENTS	83,319	91.0	9.0	55.4	35.6	236.0
SPECIAL ^c EDUCATION STUDENTS	14,508	71.9	28.1	53.1	18.8	217.2
CURRENT LIMITED ENGLISH ^d PROFICIENT STUDENTS	4,330	65.1	34.9	53.1	12.0	210.5
FORMER LIMITED ENGLISH ^d PROFICIENT STUDENTS	2,252	84.2	15.8	61.0	23.2	226.9
TOTAL LIMITED ENGLISH ^d PROFICIENT STUDENTS	6,582	71.6	28.4	55.8	15.8	216.1
TOTAL ^e STUDENTS	101,800	87.3	12.7	55.0	32.3	232.3

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
- b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
- c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
- d. INCLUDES CURRENT, FORMER, OR TOTAL (CURRENT+FORMER) LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
- e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.3.3

**NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007
 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE
 LANGUAGE ARTS LITERACY AND MATHEMATICS
 FOR THE SPECIAL NEEDS DISTRICTS AS COMPARED TO ALL OTHER DISTRICTS – Grade 3**

LANGUAGE ARTS LITERACY SECTION		NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
				PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL ^b EDUCATION STUDENTS	SPECIAL NEEDS	15,639	75.0	25.0	72.0	3.0	210.4
	ALL OTHERS	67,579	92.4	7.6	81.4	11.0	224.7
SPECIAL ^c EDUCATION STUDENTS	SPECIAL NEEDS	2,731	32.5	67.5	32.0	0.4	183.7
	ALL OTHERS	11,636	62.6	37.4	60.4	2.2	202.8
CURRENT ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	2,231	48.3	51.7	47.3	1.0	193.9
	ALL OTHERS	1,394	56.1	43.9	54.2	1.9	199.3
FORMER ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	1,111	74.3	25.7	71.6	2.7	209.4
	ALL OTHERS	1,138	80.7	19.3	76.4	4.3	214.0
TOTAL ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	3,342	57.0	43.0	55.4	1.6	199.1
	ALL OTHERS	2,231	48.3	51.7	47.3	1.0	193.9
TOTAL ^e STUDENTS	SPECIAL NEEDS	20,413	66.8	33.2	64.4	2.5	205.3
	ALL OTHERS	80,464	87.6	12.4	78.0	9.6	221.2

a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.

b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.

c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.

d. INCLUDES CURRENT, FORMER, OR TOTAL (CURRENT+FORMER) LIMITED ENGLISH PROFICIENT STUDENTS ONLY.

e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.3.3 (continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007
NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE
LANGUAGE ARTS LITERACY AND MATHEMATICS
FOR THE SPECIAL NEEDS DISTRICTS AS COMPARED TO ALL OTHER DISTRICTS – Grade 3**

MATHEMATICS SECTION		NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
				PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL ^b EDUCATION STUDENTS	SPECIAL NEEDS	15,684	78.5	21.5	59.7	18.8	221.3
	ALL OTHERS	67,635	93.9	6.1	54.5	39.5	239.4
SPECIAL ^c EDUCATION STUDENTS	SPECIAL NEEDS	2,790	50.3	49.7	42.2	8.1	199.3
	ALL OTHERS	11,718	77.0	23.0	55.6	21.4	221.5
CURRENT ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	2,543	62.4	37.6	51.9	10.5	208.2
	ALL OTHERS	1,787	68.9	31.1	54.8	14.0	213.7
FORMER ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	1,111	78.9	21.1	58.5	20.4	223.0
	ALL OTHERS	1,141	89.3	10.7	63.4	25.9	230.7
TOTAL ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	3,654	67.5	32.5	53.9	13.5	212.7
	ALL OTHERS	2,928	76.8	23.2	58.2	18.7	220.3
TOTAL ^e STUDENTS	SPECIAL NEEDS	20,815	73.0	27.0	56.5	16.5	217.0
	ALL OTHERS	80,985	91.0	9.0	54.7	36.4	236.3

a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.

b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.

c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.

d. INCLUDES CURRENT, FORMER, OR TOTAL (CURRENT+FORMER) LIMITED ENGLISH PROFICIENT STUDENTS ONLY.

e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.4.1

**NEW JERSEY STATEWIDE TESTING SYSTEMSPRING 2007
NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE BY DISTRICT
FACTOR GROUP**

LANGUAGE ARTS LITERACY SECTION – Grade 4

GENERAL EDUCATION STUDENTS^b

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	12,778	71.8	28.2	69.4	2.4	208.3
B	8,077	78.9	21.1	75.8	3.1	212.7
CD	7,782	85.2	14.8	80.8	4.4	217.1
DE	10,662	89.9	10.1	83.2	6.7	221.1
FG	10,442	91.8	8.2	84.0	7.9	222.9
GH	11,491	93.0	7.0	82.8	10.2	225.1
I	16,193	95.9	4.1	82.0	13.9	228.8
J	3,852	97.5	2.5	81.1	16.4	231.0

SPECIAL EDUCATION STUDENTS^c

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	2,700	25.7	74.3	25.4	0.3	177.5
B	1,583	38.9	61.1	38.5	0.4	188.1
CD	1,424	38.3	61.7	37.6	0.7	188.8
DE	1,956	52.5	47.5	51.9	0.6	196.5
FG	1,930	54.6	45.4	53.2	1.4	197.2
GH	1,971	57.8	42.2	56.2	1.6	199.5
I	2,952	67.5	32.5	64.9	2.6	205.5
J	535	78.7	21.3	76.4	2.2	211.0

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
d. INCLUDES CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.4.1 (continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM
 SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE
 BY DISTRICT FACTOR GROUP**

LANGUAGE ARTS LITERACY SECTION – Grade 4

CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ^d

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	1,667	41.6	58.4	41.3	0.3	188.1
B	439	38.5	61.5	38.3	0.2	187.1
CD	186	41.4	58.6	41.4	0.0	189.5
DE	200	54.5	45.5	54.0	0.5	198.2
FG	147	49.0	51.0	46.9	2.0	193.4
GH	157	51.0	49.0	49.0	1.9	195.7
I	176	65.3	34.7	63.1	2.3	204.9
J	35	68.6	31.4	60.0	8.6	206.3

TOTAL STUDENTS ^e

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	16,958	62.0	38.0	60.1	1.8	201.8
B	10,073	71.0	29.0	68.4	2.6	207.8
CD	9,373	77.3	22.7	73.5	3.8	212.4
DE	12,796	83.7	16.3	78.1	5.7	217.0
FG	12,501	85.6	14.4	78.8	6.8	218.7
GH	13,600	87.5	12.5	78.7	8.8	221.1
I	19,301	91.3	8.7	79.2	12.1	225.0
J	4,420	95.0	5.0	80.4	14.6	228.4

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
- b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
- c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
- d. INCLUDES CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
- e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.4.1 (continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM
SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE**

LANGUAGE ARTS LITERACY SECTION – Grade 4

CHARTER SCHOOLS^f

	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL EDUCATION STUDENTS ^b	1,052	67.6	32.4	64.3	3.3	206.4
SPECIAL EDUCATION STUDENTS ^c	105	29.5	70.5	29.5	0.0	181.0
CURRENT LIMITED ENGLISH ^d PROFICIENT STUDENTS	31	61.3	38.7	61.3	0.0	196.7
FORMER LIMITED ENGLISH ^d PROFICIENT STUDENTS	0	--	--	--	--	--
TOTAL LIMITED ENGLISH ^d PROFICIENT STUDENTS	31	61.3	38.7	61.3	0.0	196.7
TOTAL STUDENTS ^e	1,186	64.1	35.9	61.1	3.0	203.9

STATEWIDE RESULTS

	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL EDUCATION STUDENTS ^b	82,641	87.4	12.6	79.6	7.9	220.4
SPECIAL EDUCATION STUDENTS ^c	15,229	49.5	50.5	48.3	1.2	194.1
CURRENT LIMITED ENGLISH ^d PROFICIENT STUDENTS	3,063	44.6	55.4	43.9	0.7	190.5
FORMER LIMITED ENGLISH ^d PROFICIENT STUDENTS	1,767	68.4	31.6	65.9	2.4	206.5
TOTAL LIMITED ENGLISH ^d PROFICIENT STUDENTS	4,830	53.3	46.7	52.0	1.3	196.4
TOTAL STUDENTS ^e	100,617	80.6	19.4	73.9	6.7	215.7

a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.

b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.

c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.

d. INCLUDES CURRENT, FORMER, OR TOTAL (CURRENT+FORMER) LIMITED ENGLISH PROFICIENT STUDENTS ONLY.

e. INCLUDES ALL STUDENTS TESTED.

f. CHARTER SCHOOLS ARE NOT INCLUDED IN A DFG.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.4.2

**NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007
NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE BY DISTRICT
FACTOR GROUP**

MATHEMATICS SECTION – Grade 4**GENERAL EDUCATION STUDENTS ^b**

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	12,783	76.2	23.8	49.5	26.7	222.7
B	8,077	83.1	16.9	51.7	31.4	229.2
CD	7,786	88.3	11.7	49.9	38.4	235.0
DE	10,660	91.5	8.5	46.4	45.0	239.8
FG	10,441	93.4	6.6	46.3	47.1	241.9
GH	11,494	94.3	5.7	40.6	53.7	245.5
I	16,203	96.8	3.2	34.8	62.1	250.7
J	3,851	98.4	1.6	29.6	68.9	254.5

SPECIAL EDUCATION STUDENTS ^c

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	2,721	43.3	56.7	34.2	9.1	193.8
B	1,594	54.4	45.6	39.0	15.4	204.4
CD	1,428	56.0	44.0	41.0	15.1	204.9
DE	1,966	69.3	30.7	45.9	23.4	217.1
FG	1,933	67.4	32.6	48.1	19.3	214.0
GH	1,972	68.7	31.3	45.4	23.3	217.0
I	2,959	79.1	20.9	47.3	31.8	227.0
J	534	83.3	16.7	47.8	35.6	230.8

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
d. INCLUDES CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.4.2 (continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM
 SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE
 BY DISTRICT FACTOR GROUP**

MATHEMATICS SECTION – Grade 4

CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ^d

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	1,930	52.3	47.7	40.1	12.2	202.6
B	556	47.8	52.2	37.4	10.4	197.4
CD	230	56.5	43.5	42.6	13.9	205.6
DE	243	64.2	35.8	43.2	21.0	212.5
FG	191	58.1	41.9	39.8	18.3	208.5
GH	225	56.9	43.1	40.4	16.4	205.8
I	216	75.0	25.0	43.1	31.9	225.1
J	44	84.1	15.9	38.6	45.5	235.6

TOTAL STUDENTS ^e

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	17,232	68.7	31.3	46.2	22.5	216.3
B	10,199	76.8	23.2	49.0	27.8	223.8
CD	9,425	82.7	17.3	48.4	34.3	229.8
DE	12,846	87.6	12.4	46.3	41.3	235.9
FG	12,547	88.9	11.1	46.5	42.4	237.1
GH	13,672	90.1	9.9	41.3	48.8	240.8
I	19,356	93.9	6.1	36.7	57.2	246.8
J	4,427	96.5	3.5	31.9	64.6	251.5

a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.

b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.

c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.

d. INCLUDES CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ONLY.

e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.4.2 (continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM
SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE**

MATHEMATICS SECTION – Grade 4

CHARTER SCHOOLS^f

	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL EDUCATION STUDENTS ^b	1,052	71.0	29.0	49.0	22.1	218.4
SPECIAL EDUCATION STUDENTS ^c	105	49.5	50.5	41.0	8.6	200.7
CURRENT LIMITED ENGLISH ^d PROFICIENT STUDENTS	31	45.2	54.8	35.5	9.7	195.7
FORMER LIMITED ENGLISH ^d PROFICIENT STUDENTS	0	--	--	--	--	--
TOTAL LIMITED ENGLISH ^d PROFICIENT STUDENTS	31	45.2	54.8	35.5	9.7	195.7
TOTAL ^e STUDENTS	1,186	68.5	31.5	48.0	20.6	216.3

STATEWIDE RESULTS

	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL EDUCATION STUDENTS ^b	82,659	89.7	10.3	43.9	45.8	239.3
SPECIAL EDUCATION STUDENTS ^c	15,286	63.7	36.3	43.1	20.6	212.2
CURRENT LIMITED ENGLISH ^d PROFICIENT STUDENTS	3,701	54.8	45.2	40.2	14.6	204.7
FORMER LIMITED ENGLISH ^d PROFICIENT STUDENTS	1,764	77.7	22.3	51.1	26.6	223.7
TOTAL LIMITED ENGLISH ^d PROFICIENT STUDENTS	5,465	62.2	37.8	43.7	18.5	210.8
TOTAL ^e STUDENTS	101,310	84.7	15.3	43.7	41.0	234.1

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
- b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
- c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
- d. INCLUDES CURRENT, FORMER, OR TOTAL (CURRENT+FORMER) LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
- e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.4.3

**NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007
NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE BY DISTRICT
FACTOR GROUP**

SCIENCE SECTION – Grade 4**GENERAL EDUCATION STUDENTS^b**

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	12,775	68.6	31.4	48.9	19.8	214.7
B	8,078	80.5	19.5	49.6	30.9	225.8
CD	7,781	86.3	13.7	48.3	38.0	232.2
DE	10,652	91.3	8.7	44.9	46.3	238.2
FG	10,438	94.2	5.8	41.8	52.4	242.2
GH	11,484	93.7	6.3	39.5	54.2	243.1
I	16,185	96.3	3.7	34.3	62.0	247.9
J	3,848	98.1	1.9	29.2	68.9	251.6

SPECIAL EDUCATION STUDENTS^c

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	2,727	41.3	58.7	33.0	8.3	192.4
B	1,600	58.9	41.1	42.4	16.4	206.9
CD	1,437	60.3	39.7	43.8	16.6	208.6
DE	1,966	74.2	25.8	48.9	25.3	220.2
FG	1,932	73.2	26.8	49.6	23.6	218.1
GH	1,973	73.8	26.2	47.3	26.5	219.9
I	2,959	79.8	20.2	45.8	34.0	226.4
J	538	86.1	13.9	48.0	38.1	231.5

a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.

b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.

c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.

d. INCLUDES CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ONLY.

e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.4.3 (continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM
 SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE
 BY DISTRICT FACTOR GROUP**

SCIENCE SECTION – Grade 4

CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ^d

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	1,925	38.4	61.6	30.9	7.5	190.5
B	555	35.5	64.5	30.1	5.4	187.3
CD	227	36.1	63.9	29.1	7.0	191.1
DE	245	52.7	47.3	36.3	16.3	204.0
FG	190	46.8	53.2	35.3	11.6	198.3
GH	224	42.9	57.1	34.4	8.5	194.0
I	214	64.5	35.5	45.3	19.2	210.5
J	43	81.4	18.6	69.8	11.6	218.0

TOTAL STUDENTS ^e

DFG	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
A	17,226	61.4	38.6	44.6	16.7	208.8
B	10,205	74.8	25.2	47.5	27.3	220.9
CD	9,426	81.3	18.7	47.2	34.1	227.7
DE	12,840	88.0	12.0	45.4	42.6	234.9
FG	12,542	90.3	9.7	42.9	47.4	237.9
GH	13,662	90.0	10.0	40.5	49.5	239.0
I	19,336	93.5	6.5	36.2	57.3	244.2
J	4,427	96.5	3.5	31.8	64.6	248.9

a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.

b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.

c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.

d. INCLUDES CURRENT LIMITED ENGLISH PROFICIENT STUDENTS ONLY.

e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.4.3 (continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM
SPRING 2007 NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE**

SCIENCE SECTION – Grade 4

CHARTER SCHOOLS^f

	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL EDUCATION STUDENTS ^b	1,048	67.7	32.3	46.3	21.5	215.1
SPECIAL EDUCATION STUDENTS ^c	105	43.8	56.2	33.3	10.5	195.4
CURRENT LIMITED ENGLISH ^d PROFICIENT STUDENTS	31	41.9	58.1	35.5	6.5	196.3
FORMER LIMITED ENGLISH ^d PROFICIENT STUDENTS	0	--	--	--	--	--
TOTAL LIMITED ENGLISH ^d PROFICIENT STUDENTS	31	41.9	58.1	35.5	6.5	196.3
TOTAL STUDENTS ^e	1,182	65.1	34.9	44.9	20.1	213.0

STATEWIDE RESULTS

	NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
			PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL EDUCATION STUDENTS ^b	82,601	87.9	12.1	42.4	45.5	236.1
SPECIAL EDUCATION STUDENTS ^c	15,311	66.4	33.6	44.0	22.4	213.8
CURRENT LIMITED ENGLISH ^d PROFICIENT STUDENTS	3,689	41.3	58.7	32.6	8.7	193.0
FORMER LIMITED ENGLISH ^d PROFICIENT STUDENTS	1,765	69.4	30.6	50.8	18.6	213.6
TOTAL LIMITED ENGLISH ^d PROFICIENT STUDENTS	5,454	50.4	49.6	38.5	11.9	199.6
TOTAL STUDENTS ^e	101,266	83.1	16.9	42.3	40.8	231.3

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
- b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
- c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
- d. INCLUDES CURRENT, FORMER, OR TOTAL (CURRENT+FORMER) LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
- e. INCLUDES ALL STUDENTS TESTED.
- f. CHARTER SCHOOLS ARE NOT INCLUDED IN A DFG.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.4.4

**NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007
NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE
LANGUAGE ARTS LITERACY, MATHEMATICS, AND SCIENCE
FOR THE SPECIAL NEEDS DISTRICTS AS COMPARED TO ALL OTHER DISTRICTS – Grade 4**

LANGUAGE ARTS LITERACY SECTION		NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
				PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL ^b EDUCATION STUDENTS	SPECIAL NEEDS	14,998	71.8	28.2	69.4	2.4	208.5
	ALL OTHERS	67,643	90.9	9.1	81.8	9.1	223.1
SPECIAL ^c EDUCATION STUDENTS	SPECIAL NEEDS	3,089	26.2	73.8	25.9	0.3	178.3
	ALL OTHERS	12,140	55.4	44.6	54.0	1.4	198.1
CURRENT ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	1,834	40.9	59.1	40.7	0.3	187.7
	ALL OTHERS	1,229	50.0	50.0	48.7	1.2	194.7
FORMER ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	939	64.6	35.4	62.3	2.3	204.2
	ALL OTHERS	828	72.6	27.4	70.0	2.5	209.1
TOTAL ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	2,773	49.0	51.0	48.0	1.0	193.3
	ALL OTHERS	2,057	59.1	40.9	57.3	1.8	200.5
TOTAL ^e STUDENTS	SPECIAL NEEDS	19,726	62.3	37.7	60.3	1.9	202.2
	ALL OTHERS	80,891	85.1	14.9	77.2	7.8	219.0

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
- b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
- c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
- d. INCLUDES CURRENT, FORMER, OR TOTAL (CURRENT+FORMER) LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
- e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.4.4 (continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007
NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE
LANGUAGE ARTS LITERACY, MATHEMATICS, AND SCIENCE
FOR THE SPECIAL NEEDS DISTRICTS AS COMPARED TO ALL OTHER DISTRICTS – Grade 4**

MATHEMATICS SECTION		NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
				PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL ^b EDUCATION STUDENTS	SPECIAL NEEDS	15,004	76.8	23.2	49.9	27.0	223.4
	ALL OTHERS	67,655	92.6	7.4	42.6	50.0	242.8
SPECIAL ^c EDUCATION STUDENTS	SPECIAL NEEDS	3,120	42.4	57.6	33.7	8.7	193.3
	ALL OTHERS	12,166	69.2	30.8	45.5	23.7	217.1
CURRENT ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	2,132	52.4	47.6	40.4	12.1	202.5
	ALL OTHERS	1,569	58.1	41.9	40.0	18.1	207.8
FORMER ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	938	74.5	25.5	49.6	24.9	221.2
	ALL OTHERS	826	81.4	18.6	52.9	28.5	226.5
TOTAL ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	3,070	59.2	40.8	43.2	16.0	208.2
	ALL OTHERS	2,395	66.1	33.9	44.4	21.7	214.3
TOTAL ^e STUDENTS	SPECIAL NEEDS	20,045	69.3	30.7	46.5	22.8	216.8
	ALL OTHERS	81,265	88.5	11.5	43.0	45.5	238.3

- a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.
- b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.
- c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.
- d. INCLUDES CURRENT, FORMER, OR TOTAL (CURRENT+FORMER) LIMITED ENGLISH PROFICIENT STUDENTS ONLY.
- e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

TABLE B.4.4 (Continued)

**NEW JERSEY STATEWIDE TESTING SYSTEM SPRING 2007
NEW JERSEY ASSESSMENT OF SKILLS AND KNOWLEDGE
LANGUAGE ARTS LITERACY, MATHEMATICS, AND SCIENCE
FOR THE SPECIAL NEEDS DISTRICTS AS COMPARED TO ALL OTHER DISTRICTS – Grade 4**

SCIENCE SECTION		NUMBER ^a TESTED 2007	PERCENT WHO SCORED ADVANCED PROFICIENT OR PROFICIENT 2007	PERCENT AT EACH PROFICIENCY LEVELS			MEAN SCALE SCORE 2007
				PARTIALLY PROFICIENT (100-199)	PROFICIENT (200-249)	ADVANCED PROFICIENT (250-300)	
GENERAL ^b EDUCATION STUDENTS	SPECIAL NEEDS	15,002	69.5	30.5	48.8	20.7	215.5
	ALL OTHERS	67,599	92.0	8.0	41.0	51.0	240.7
SPECIAL ^c EDUCATION STUDENTS	SPECIAL NEEDS	3,122	41.5	58.5	33.5	8.1	192.6
	ALL OTHERS	12,189	72.8	27.2	46.7	26.1	219.2
CURRENT ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	2,124	38.0	62.0	31.0	7.0	190.1
	ALL OTHERS	1,565	45.8	54.2	34.8	11.0	196.8
FORMER ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	938	62.8	37.2	48.2	14.6	208.5
	ALL OTHERS	827	76.9	23.1	53.8	23.1	219.4
TOTAL ^d LIMITED ENGLISH PROFICIENT STUDENTS	SPECIAL NEEDS	3,062	45.6	54.4	36.3	9.3	195.7
	ALL OTHERS	2,392	56.6	43.4	41.4	15.2	204.7
TOTAL ^e STUDENTS	SPECIAL NEEDS	20,038	62.2	37.8	44.7	17.4	209.5
	ALL OTHERS	81,228	88.3	11.7	41.8	46.6	236.7

a. EXCLUDES STUDENTS' TEST BOOKLETS CODED VOID, NOT PRESENT AND APA EXEMPT WITH NO SCALED SCORES.

b. EXCLUDES SPECIAL EDUCATION AND LIMITED ENGLISH PROFICIENT STUDENTS.

c. INCLUDES SPECIAL EDUCATION STUDENTS ONLY.

d. INCLUDES CURRENT, FORMER, OR TOTAL (CURRENT+FORMER) LIMITED ENGLISH PROFICIENT STUDENTS ONLY.

e. INCLUDES ALL STUDENTS TESTED.

NOTE: PERCENTAGES MAY NOT TOTAL 100 DUE TO ROUNDING

How to Interpret the Categories

The following is an explanation of how to interpret the categories of students presented in the following report. Please apply these rules as you read and interpret the report.

For each content area:

“General Education” excludes students coded as special education OR limited English proficient on their test booklets.

“Special Education” includes students coded as SE on their test booklet

“Limited English Proficient, Current and Former” or **“Total Limited English Proficient”** includes students coded as LEP or Former LEP on their test booklet.

“Current Limited English Proficient” includes students coded as LEP on their test booklet.

“Former Limited English Proficient” includes students coded as Former LEP on their test booklet.

“Total” includes all students tested who were not Void.

DISTRICT FACTOR GROUPS

The District Factor Group (DFG) is an indicator of the socioeconomic status of citizens in each district and has been useful for the comparative reporting of test results from New Jersey's statewide testing programs. The measure was first developed in 1974 using demographic variables from the 1970 United States Census. Revisions were made in 1984 and 1992 to take into account new data from the 1980 and 1990 United States Census. The DFG designations were updated again in 2004, using the following demographic variables from the 2000 United States Census.

- A. Percentage of adult residents who failed to complete high school
- B. Percentage of adult residents who attended college
- C. Occupational status of adult household members:
 - 1 = laborers
 - 2 = service workers (except private and protective)
 - 3 = farm workers
 - 4 = operatives and kindred workers
 - 5 = protective service workers
 - 6 = sales workers
 - 7 = clerical and kindred workers
 - 8 = craftsmen, foremen, and kindred workers
 - 9 = quasi-professionals
 - 10 = managers, officials, and proprietors
 - 11 = old and new professionals
- D. Population Density: persons per square mile
- E. Income: median family income
- F. Unemployment: percentage of those in the work force who received some unemployment compensation
- G. Poverty: percentage of residents below the poverty level

The variables described above were combined using a statistical technique called principal components analysis, which resulted in a single measure of socioeconomic status for each district. Districts were then ranked according to their score on this measure and divided into eight groups based on the score interval in which their scores were located. Eight DFGs have been created based on the 1990 United States Census data. They range from A (lowest socioeconomic districts) to J (highest socioeconomic districts) and are labeled as follows: A, B, CD, DE, FG, GH, I, J. Updating the DFGs has not changed any district's designation as Special Needs or not Special Needs.

Whereas the DFGs based on the 1970 and 1980 United States Census resulted in 10 groups containing approximately equal numbers of districts, the DFGs based on the 1990 and 2000 United States Census resulted in eight groups of different sizes depending on their score. The number of districts* in each DFG is now as follows:

DFG	Number of Districts
A	39
B	67
CD	67
DE	83
FG	89
GH	76
I	103
J	25

In the most recent DFG designations, 15 non-functioning or cooperative school districts are no longer included in a DFG. Two other districts, Deal and Lakewood, no longer carry a DFG designation because more than 50% of their students attend private schools.

* Includes all New Jersey's public school districts (regardless of school configuration or grade levels served).

**APPENDIX C:
Raw to Scale Score Conversions**

Raw Score – Scale Score Conversions with Theta, S.E. and Cumulative Frequencies

TABLE C.3.1 Conversion 2007 NJ ASK Language Arts Literacy – Grade 3

Raw Score	Scale Score	Theta	S.E.	Cumulative Number of Students *	Cumulative Percent of Students
0	128	-2.7534	1.285	8	0.0
1	132	-1.8343	0.481	50	0.0
2	136	-1.5224	0.331	123	0.1
3	140	-1.3453	0.270	275	0.3
4	144	-1.2172	0.238	467	0.5
5	148	-1.1127	0.220	761	0.8
6	152	-1.0213	0.208	1,102	1.1
7	157	-0.9377	0.201	1,555	1.5
8	161	-0.8588	0.196	2,108	2.1
9	165	-0.7824	0.194	2,797	2.8
10	169	-0.7067	0.194	3,706	3.7
11	173	-0.6300	0.197	4,907	4.9
12	177	-0.5509	0.200	6,366	6.3
13	181	-0.4684	0.205	8,248	8.2
14	186	-0.3817	0.211	10,570	10.5
15	190	-0.2903	0.216	13,250	13.1
16	194	-0.1938	0.222	16,758	16.6
17	200	-0.0922	0.228	21,159	21.0
18	203	0.0150	0.234	26,001	25.8
19	207	0.1283	0.241	31,924	31.7
20	212	0.2481	0.248	38,606	38.3

Raw Score	Scale Score	Theta	S.E.	Cumulative Number of Students *	Cumulative Percent of Students
21	217	0.3750	0.255	46,585	46.2
22	221	0.5092	0.262	55,474	55.0
23	226	0.6502	0.268	64,989	64.4
24	231	0.7977	0.274	74,152	73.5
25	235	0.9514	0.280	82,091	81.4
26	240	1.1115	0.286	87,797	87.1
27	244	1.2788	0.292	92,639	91.9
28	250	1.4529	0.297	96,543	95.7
29	252	1.6312	0.299	98,614	97.8
30	257	1.8089	0.296	99,778	98.9
31	261	1.9820	0.291	100,356	99.5
32	265	2.1495	0.287	100,647	99.8
33	269	2.3134	0.285	100,762	99.9
34	272	2.4773	0.287	100,818	100.0
35	276	2.6463	0.294	100,841	100.0
36	279	2.8284	0.310	100,852	100.0
37	281	3.0386	0.341	100,855	100.0
38	283	3.3103	0.402	100,857	100.0
39	286	3.7437	0.549	100,857	100.0
40	288	4.7974	1.318	100,857	100.0

* Excludes Students Who Took Breach Form.

TABLE C.3.2 Conversion 2007 NJ ASK Mathematics – Grade 3

Raw Score	Scale Score	Theta	S.E.	Cumulative Number of Students *	Cumulative Percent of Students
0.0	119	-3.8106	1.824	3	0.0
0.5	122	-2.6282	0.975	5	0.0
1.0	125	-1.9910	0.664	10	0.0
1.5	127	-1.6420	0.529	11	0.0
2.0	130	-1.4037	0.452	34	0.0
2.5	133	-1.2223	0.402	69	0.1
3.0	137	-1.0751	0.366	122	0.1
3.5	140	-0.9506	0.340	203	0.2
4.0	143	-0.8420	0.319	325	0.3
4.5	146	-0.7453	0.302	471	0.5
5.0	149	-0.6578	0.289	703	0.7
5.5	153	-0.5774	0.277	1,008	1.0
6.0	156	-0.5030	0.268	1,414	1.4
6.5	159	-0.4334	0.259	1,864	1.8
7.0	162	-0.3679	0.252	2,397	2.4
7.5	166	-0.3059	0.245	3,009	3.0
8.0	169	-0.2470	0.240	3,723	3.7
8.5	172	-0.1907	0.234	4,509	4.4
9.0	175	-0.1367	0.230	5,362	5.3
9.5	178	-0.0848	0.225	6,247	6.1
10.0	181	-0.0347	0.221	7,217	7.1
10.5	184	0.0137	0.218	8,195	8.1
11.0	187	0.0604	0.214	9,261	9.1
11.5	190	0.1058	0.211	10,306	10.1
12.0	193	0.1499	0.208	11,561	11.4
12.5	196	0.1927	0.205	12,843	12.6
13.0	200	0.2344	0.202	14,356	14.1
13.5	201	0.2750	0.200	15,801	15.5
14.0	204	0.3145	0.197	17,444	17.2
14.5	206	0.3532	0.195	18,924	18.6
15.0	209	0.3909	0.193	20,691	20.4
15.5	211	0.4279	0.191	22,306	21.9
16.0	214	0.4641	0.189	24,201	23.8
16.5	216	0.4996	0.187	25,955	25.5
17.0	218	0.5345	0.186	28,153	27.7
17.5	220	0.5689	0.184	29,881	29.4
18.0	222	0.6028	0.183	32,307	31.8
18.5	224	0.6364	0.182	34,191	33.6
19.0	226	0.6697	0.182	36,839	36.2
19.5	228	0.7029	0.181	38,775	38.1
20.0	230	0.7359	0.181	41,601	40.9
20.5	232	0.7690	0.182	43,563	42.9
21.0	234	0.8022	0.182	46,754	46.0
21.5	236	0.8357	0.183	48,728	47.9
22.0	238	0.8696	0.184	52,126	51.3
22.5	240	0.9040	0.186	54,199	53.3
23.0	241	0.9390	0.188	57,867	56.9
23.5	243	0.9749	0.190	59,598	58.6
24.0	245	1.0119	0.193	63,072	62.0
24.5	247	1.0502	0.197	64,868	63.8
25.0	249	1.0899	0.201	68,804	67.7
25.5	250	1.1314	0.206	70,749	69.6
26.0	252	1.1750	0.211	75,238	74.0
26.5	254	1.2212	0.218	76,807	75.6
27.0	256	1.2704	0.225	80,986	79.7
27.5	257	1.3233	0.234	82,353	81.0
28.0	259	1.3806	0.244	86,415	85.0
28.5	261	1.4434	0.256	87,485	86.1
29.0	262	1.5130	0.271	91,330	89.8
29.5	264	1.5913	0.289	92,260	90.8
30.0	266	1.6810	0.311	95,574	94.0
30.5	267	1.7865	0.339	96,136	94.6
31.0	269	1.9146	0.378	98,688	97.1
31.5	271	2.0785	0.435	98,985	97.4
32.0	272	2.3083	0.532	100,763	99.1
32.5	274	2.7021	0.756	100,878	99.2
33.0	276	3.4110	1.427	101,654	100.0

* Excludes Students Who Took Breach Form and Large Print Form.

TABLE C.4.1 Conversion 2007 NJ ASK Language Arts Literacy – Grade 4

Raw Score	Scale Score	Theta	S.E.	Cumulative Number of Students *	Cumulative Percent of Students	Raw Score	Scale Score	Theta	S.E.	Cumulative Number of Students *	Cumulative Percent of Students
0	107	-2.6804	1.275	5	0.0	22	216	0.3509	0.231	48,003	47.7
1	112	-1.7998	0.461	34	0.0	23	220	0.4605	0.236	55,640	55.3
2	117	-1.5164	0.314	94	0.1	24	225	0.5753	0.242	63,878	63.5
3	122	-1.3554	0.259	169	0.2	25	229	0.6959	0.248	71,789	71.4
4	128	-1.2361	0.232	314	0.3	26	233	0.8228	0.255	79,155	78.7
5	133	-1.1358	0.217	485	0.5	27	237	0.9558	0.260	85,463	84.9
6	140	-1.0458	0.207	775	0.8	28	241	1.0936	0.263	90,180	89.6
7	146	-0.9625	0.200	1,175	1.2	29	244	1.2334	0.264	93,890	93.3
8	152	-0.8840	0.195	1,748	1.7	30	250	1.3720	0.261	96,784	96.2
9	157	-0.8086	0.192	2,493	2.5	31	251	1.5068	0.257	98,403	97.8
10	162	-0.7349	0.191	3,448	3.4	32	255	1.6369	0.252	99,392	98.8
11	167	-0.6609	0.193	4,652	4.6	33	258	1.7633	0.250	99,944	99.3
12	171	-0.5850	0.196	6,186	6.1	34	261	1.8880	0.249	100,289	99.7
13	175	-0.5057	0.201	8,058	8.0	35	264	2.0141	0.253	100,446	99.8
14	180	-0.4223	0.206	10,321	10.3	36	267	2.1454	0.260	100,545	99.9
15	184	-0.3348	0.211	13,049	13.0	37	270	2.2872	0.273	100,587	100.0
16	189	-0.2438	0.215	15,927	15.8	38	274	2.4469	0.293	100,604	100.0
17	194	-0.1501	0.217	19,524	19.4	39	277	2.6352	0.322	100,608	100.0
18	200	-0.0543	0.219	23,965	23.8	40	281	2.8697	0.364	100,611	100.0
19	203	0.0433	0.222	29,091	28.9	41	285	3.1819	0.429	100,611	100.0
20	207	0.1430	0.224	34,568	34.4	42	289	3.6576	0.564	100,611	100.0
21	212	0.2453	0.227	40,759	40.5	43	293	4.7265	1.318	100,611	100.0

* Excludes Students Who Took Breach Form.

TABLE C.4.2 Conversion 2007 NJ ASK Mathematics – Grade 4

Raw Score	Scale Score	Theta	S.E.	Cumulative Number of Students *	Cumulative Percent of Students	Raw Score	Scale Score	Theta	S.E.	Cumulative Number of Students *	Cumulative Percent of Students
0.0	104	-4.2007	1.862	0	0.0	22.0	222	0.7848	0.154	33,241	32.8
0.5	107	-2.9144	1.047	0	0.0	22.5	224	0.8084	0.153	34,688	34.2
1.0	110	-2.1496	0.740	3	0.0	23.0	226	0.8319	0.153	36,574	36.1
1.5	114	-1.7135	0.591	8	0.0	23.5	227	0.8553	0.152	37,941	37.5
2.0	118	-1.4191	0.499	32	0.0	24.0	229	0.8786	0.152	40,043	39.5
2.5	122	-1.2013	0.437	53	0.1	24.5	231	0.9019	0.152	41,454	40.9
3.0	127	-1.0299	0.392	105	0.1	25.0	233	0.9252	0.152	43,582	43.0
3.5	131	-0.8890	0.359	174	0.2	25.5	234	0.9485	0.153	44,971	44.4
4.0	134	-0.7693	0.333	283	0.3	26.0	236	0.9720	0.153	47,226	46.6
4.5	138	-0.6650	0.312	415	0.4	26.5	238	0.9956	0.154	48,644	48.0
5.0	142	-0.5727	0.295	642	0.6	27.0	240	1.0195	0.154	51,014	50.4
5.5	145	-0.4895	0.281	887	0.9	27.5	241	1.0435	0.155	52,344	51.7
6.0	148	-0.4137	0.269	1,242	1.2	28.0	243	1.0679	0.156	54,628	53.9
6.5	151	-0.3439	0.259	1,616	1.6	28.5	245	1.0926	0.157	55,900	55.2
7.0	155	-0.2792	0.249	2,081	2.1	29.0	246	1.1177	0.159	58,448	57.7
7.5	158	-0.2187	0.241	2,577	2.5	29.5	248	1.1433	0.160	59,760	59.0
8.0	160	-0.1620	0.234	3,144	3.1	30.0	250	1.1695	0.162	62,687	61.9
8.5	163	-0.1084	0.228	3,805	3.8	30.5	251	1.1962	0.164	64,117	63.3
9.0	166	-0.0576	0.222	4,415	4.4	31.0	253	1.2236	0.166	66,837	66.0
9.5	169	-0.0093	0.217	5,058	5.0	31.5	254	1.2517	0.169	68,101	67.2
10.0	171	0.0367	0.212	5,767	5.7	32.0	256	1.2807	0.171	70,887	70.0
10.5	174	0.0808	0.207	6,514	6.4	32.5	257	1.3107	0.174	72,066	71.1
11.0	176	0.1230	0.203	7,272	7.2	33.0	259	1.3417	0.177	74,911	73.9
11.5	179	0.1636	0.199	8,083	8.0	33.5	260	1.3739	0.181	76,040	75.1
12.0	181	0.2026	0.195	8,953	8.8	34.0	262	1.4074	0.185	78,905	77.9
12.5	183	0.2402	0.192	9,752	9.6	34.5	263	1.4425	0.189	79,929	78.9
13.0	186	0.2765	0.188	10,695	10.6	35.0	265	1.4792	0.194	82,758	81.7
13.5	188	0.3116	0.185	11,607	11.5	35.5	266	1.5180	0.199	83,674	82.6
14.0	190	0.3456	0.182	12,556	12.4	36.0	267	1.5589	0.205	86,555	85.4
14.5	192	0.3784	0.179	13,464	13.3	36.5	269	1.6026	0.212	87,351	86.2
15.0	194	0.4103	0.177	14,492	14.3	37.0	270	1.6492	0.220	89,993	88.8
15.5	196	0.4413	0.174	15,524	15.3	37.5	271	1.6995	0.228	90,700	89.5
16.0	200	0.4714	0.172	16,778	16.6	38.0	272	1.7542	0.239	93,280	92.1
16.5	201	0.5007	0.170	17,903	17.7	38.5	273	1.8142	0.251	93,815	92.6
17.0	203	0.5292	0.167	19,173	18.9	39.0	274	1.8808	0.265	96,062	94.8
17.5	205	0.5571	0.165	20,335	20.1	39.5	275	1.9558	0.282	96,458	95.2
18.0	207	0.5843	0.164	21,670	21.4	40.0	277	2.0419	0.304	98,268	97.0
18.5	209	0.6109	0.162	22,853	22.6	40.5	278	2.1431	0.332	98,526	97.3
19.0	211	0.6369	0.160	24,290	24.0	41.0	279	2.2665	0.371	99,943	98.7
19.5	213	0.6625	0.159	25,546	25.2	41.5	280	2.4251	0.428	100,071	98.8
20.0	214	0.6877	0.157	27,158	26.8	42.0	281	2.6490	0.526	100,945	99.6
20.5	216	0.7124	0.156	28,419	28.1	42.5	282	3.0362	0.751	100,995	99.7
21.0	218	0.7368	0.155	30,116	29.7	43.0	285	3.7398	1.424	101,304	100.0
21.5	220	0.7610	0.154	31,456	31.1						

* Excludes Students Who Took Breach Form.

TABLE C.4.3 Conversion 2007 NJ ASK Science – Grade 4

Raw Score	Scale Score	Theta	S.E.	Cumulative Number of Students *	Cumulative Percent of Students
0	114	-2.8116	1.294	2	0.0
1	119	-1.8457	0.510	9	0.0
2	124	-1.4784	0.368	18	0.0
3	129	-1.2546	0.306	47	0.0
4	134	-1.0895	0.270	123	0.1
5	139	-0.9561	0.246	284	0.3
6	144	-0.8429	0.229	572	0.6
7	148	-0.7433	0.217	1,018	1.0
8	153	-0.6536	0.207	1,587	1.6
9	158	-0.5713	0.199	2,411	2.4
10	162	-0.4946	0.192	3,439	3.4
11	167	-0.4225	0.187	4,601	4.5
12	171	-0.3539	0.183	5,970	5.9
13	176	-0.2881	0.179	7,664	7.6
14	181	-0.2245	0.176	9,673	9.6
15	185	-0.1627	0.174	11,936	11.8
16	190	-0.1022	0.173	14,264	14.1
17	195	-0.0427	0.172	17,071	16.9
18	200	0.0162	0.171	20,503	20.2
19	205	0.0749	0.171	23,978	23.7
20	209	0.1335	0.171	27,753	27.4
21	214	0.1926	0.172	31,701	31.3
22	219	0.2523	0.173	35,983	35.5
23	224	0.3131	0.175	40,501	40.0
24	229	0.3753	0.177	45,200	44.6
25	234	0.4393	0.180	50,121	49.5
26	238	0.5056	0.183	54,681	54.0
27	243	0.5748	0.188	59,953	59.2
28	250	0.6475	0.193	66,016	65.2
29	252	0.7246	0.199	71,706	70.8
30	256	0.8071	0.206	77,243	76.3
31	260	0.8964	0.216	82,560	81.5
32	264	0.9944	0.227	87,457	86.4
33	268	1.1039	0.241	91,867	90.7
34	272	1.2287	0.259	95,498	94.3
35	276	1.3758	0.284	98,196	97.0
36	280	1.5572	0.320	99,920	98.7
37	283	1.8001	0.381	100,890	99.6
38	287	2.1895	0.521	101,204	99.9
39	291	3.1768	1.300	101,260	100.0

* Excludes Students Who Took Breach Form.

REFERENCES

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for Educational and Psychological Testing*. Washington, DC: Author.
- Cizek, G. J. (1993). *Reactions to National Academy of Education report, Setting performance standards for student achievement*. Washington, DC: National Assessment Governing Board.
- Directory of Test Specifications and Sample Items for the Elementary School Proficiency Assessment (ESPA) in Mathematics*. Developed by the New Jersey Department of Education, April 1998.
- Directory of Test Specifications and Sample Items for the Elementary School Proficiency Assessment (ESPA) in Science*. Developed by the New Jersey Department of Education, April 1998.
- Directory of Test Specifications and Sample Items for the Elementary School Proficiency Assessment (ESPA), Grade Eight Proficiency Assessment (GEPA), and High School Proficiency Assessment (HSPA) in Language Arts Literacy*. Developed by the New Jersey Department of Education, February 1998.
- Directory of Test Specifications and Sample Items for the Grade Eight Proficiency Assessment (GEPA) and the High School Proficiency Assessment (HSPA) in Mathematics*. Developed for the New Jersey Department of Education, February 1998.
- Elementary School Proficiency Assessment (ESPA) Standard Setting Summary. 1999 Language Arts Literacy, Mathematics and Science*. Developed for the New Jersey Department of Education by National Computer Systems, August 1999.
- Fitzpatrick, A.R., Ercikan, K., Yen, W.M., and Ferrara, S. (1998). The consistency between raters scoring in different test years. *Applied Measurement in Education*, 11(2), 195-208.
- Grade Eight Proficiency Assessment (GEPA) Standard Setting Study Report of Activities. 1999 Language Arts Literacy and Mathematics, and 2000 Science*. Developed for the New Jersey Department of Education by National Computer Systems, November 2000.
- Hambleton, R. K. & Plake, B. S. (1995). Using an extended Angoff procedure to set standards on complex performance assessments. *Applied Measurement in Education*, 8, 41-55.
- Hambleton, R., Swaminathan, H., & Rogers, H. (1991). *Fundamentals of Item Response Theory*. Newbury Park: Sage Publications, Inc.

- Holland, P. W. & Thayer, D. T. (1986). *Differential Item Functioning and the Mantel-Haenszel Procedure* (Technical Rep. No. 86-69). Princeton, NJ: Educational Testing Service.
- Kane, M. (1995). Examinee-centered vs. task-centered standard setting. In *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments* (pp. 119-139). Washington, DC: National Assessment Governing Board and National Center for Education Statistics.
- Kolen, M. J. & Brennan, R. L. (1995). *Test Equating*. New York: Springer-Verlag.
- Livingston, S. A., and Lewis, C. (1995). Estimating the consistency and accuracy of classification based on test scores. *Journal on Educational Measurement*, 32, 179-197.
- Loomis, S. C. & Borque, M. L. (2001). From tradition to innovation: Standard setting on the National Assessment of Educational Progress. In G. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 175-217). Mahwah, NJ: Lawrence Erlbaum.
- Mehrens, W. A. (1995). Methodological issues in standard setting for educational exams. In *Proceedings of Joint Conference on Standard Setting for Large-Scale Assessments* (pp. 221-263). Washington, DC: National Assessment Governing Board and National Center for Education Statistics.
- Mehrens, W. A. & Lehmann, I. J. (1991). *Measurement and Evaluation in Education and Psychology* (4th ed.). New York: Holt, Rinehart and Winston.
- Millman, J., & Greene, J. (1989). The specification and development of tests of achievement and ability. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 335-366). New York: Macmillan.
- National Academy of Education (1993). *Setting performance standards for student achievement*. Washington, DC.
- SAS Institute (1985). *User's Guide: Statistics* (Vol. I and II). Cary, NC: Author.